OCT 19 1928

PERIODICAL ROOM GENERAL LIBRARY UNIV. OF MICH.

The Journal of EDUCATIONAL PSYCHOLOGY

Devoted Primarily to the Scientific Study of Problems of Learning and Teaching.

BOARD OF EDITORS

HAROLD RUGG, *Chairman.*

J. CARLETON BELL — ARTHUR I. GATES — RUDOLF PINTNER — L. M. TERMAN

FRANK N. FREEMAN — V. A. C. HENMON — BEARDSLEY RUML — E. L. THORNDIKE

VOL. XIX
No. 7

OCTOBER, 1928

$5.00 a year.
70c per copy.

CONTENTS

SCALE CONSTRUCTION WITH WEIGHTED OBSERVATIONS. *L. L. Thurstone* 441

LECTURE AND CLASS DISCUSSION IN TEACHING EDUCATIONAL PSYCHOLOGY. *Ralph B. Spence* 454

PRESENT STATUS OF THE QUESTION OF RACIAL DIFFERENCES. *Dale Yoder* 463

THE CURVE OF RETENTION FOR POETRY. *Paul L. Whitely and John A. McGeoch* 471

THE RELATIVE INFLUENCE OF TWO TYPES OF MOTIVATION ON IMPROVEMENT. *Verner Martin Sims* 480

LEARNING AND RETENTION OF LATIN WORDS AND PHRASES. *J. P. Anderson and A. M. Jordan* 485

TRUMAN L. KELLEY MEASURES MENTAL TRAITS. *C. S. Slocombe* 497

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES 502

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION 508

PUBLICATIONS RECEIVED 516

Published Monthly Except June to August by

Warwick & York, Inc.,

York, Pa. Baltimore, Md.

Entered as Second Class matter November 15, 1921, at the Post Office at York, Pennsylvania under the Act of March 3, 1879.

The Journal of Educational Psychology

is published monthly, by Warwick & York, Inc., Baltimore, Md., and York, Pa. Title page and index are bound in December number.

Manuscripts for publication, books or other materials for review, and news items should be addressed to Harold Rugg, Chairman of the Editorial Board, 400 W. 118th Street, New York City.

The price of the Journal is $5.00 a year, payable in advance. Foreign postage is 30 cents extra. Single current issues are 70 cents each, and less than a full year's subscription is charged at single issue rate for every month ordered. Back volumes $5.50 each; single back issues 90 cents each.

Subscribers should notify the publishers of change in address at least four weeks before publication of the issue with which change is to take effect. No claim for nonreceipt of an issue will be entertained unless made within two weeks after receipt of the next succeeding issue.

Warwick & York, Inc.,

Baltimore, Md.

The Commonwealth Fund Teacher Training Study

By WERRITT W. CHARTERS
and DOUGLAS WAPLES

Materials for Developing Teacher Training Curricula and Methods of Supervision.

To be published in February

The UNIVERSITY *of* CHICAGO PRESS
CHICAGO - - - ILLINOIS

TEACHERS' HEADQUARTERS

European
Plan

Entirely
Fire-Proof

Centrally
Located

EDWARD DAVIS
Manager

HOTEL RENNERT
BALTIMORE, MARYLAND

The Value of the Dictionary

is the title of a *new* booklet written to help you in Teaching the Dictionary. Here are a few suggestions of the lessons included:

First Dictionary Lessons
Relative Position of Letters
How to Find Words
What You Find
Pronunciation
How to Find Meanings
Parts of Speech and Meanings
Unusual Uses of Words
Synonyms
The Hyphen, Etc., Etc.

Copies of this new booklet will be sent FREE to teachers upon request.

G. & C. Merriam Company
Springfield, Mass.

Publishers of the Merriam-Webster Dictionaries for over 85 years
Look for the Circular Trade-Mark Get the Best

WEBSTER'S NEW INTERNATIONAL DICTIONARY

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume XIX October, 1928 Number 7

SCALE CONSTRUCTION WITH WEIGHTED OBSERVATIONS

L. L. THURSTONE

University of Chicago

In a previous article in this journal, I described what I have called an absolute scaling method.[2] It is called an absolute scaling method because it is independent of the scoring unit represented by the raw scores. The unit of measurement is the standard deviation of the distribution of test ability in any given age group. This distribution of ability is assumed, as usual, to be normal and the base line is an abstract scale of test ability independent of the raw scoring unit. The method previously described enables one to represent a series of overlapping age or grade groups on the same scale.

It is the purpose of this paper to develop the same scaling technique in such a way that the observed proportion of correct answers, or the observed proportion of each age group that exceeds a given raw score, may be weighted in accordance with its probable error. As an example of the method, we shall apply it to the data of Woody's arithmetic scales.[3]

Our object is to represent the seven overlapping grade groups described by Woody on a single scale of test ability. This is shown in our Fig. 1. The overlapping grade distributions have been drawn in Fig. 1 on separate base lines in order to make clearer their relative dispersions. The arbitrary origin for the absolute scale of Fig. 1 is

[1] This is one of a series of studies by the staff of the Behavior Research Fund of the Illinois Institute for Juvenile Research, Dr. Herman M. Adler, director. Series B No. 111.

[2] Thurstone, L. L.: A Method of Scaling Psychological and Educational Tests. *The Journal of Educational Psychology*, Vol. XVI, Oct., 1925, pp. 433–451.

[3] Woody, C.: "Measurements of Some Achievements in Arithmetic." Teachers College, Columbia University Contributions to Education, No. 80.

placed at the mean test ability of second grade children and the standard deviation of that grade is chosen as a unit of measurement for the absolute scale. Each of the test items is located at a point on the common scale. If the data were free from errors of measurement the proportion of any grade distribution to the right of that point would represent the proportion of children in that grade who answer the test item correctly. This should be true for all grades whose distributions cover the scale value of the test item.

In practice, such perfection is only approximately attained because of errors of measurement in the original observed proportions of right answers for each test item in each grade and also because of certain peculiarities in the drill on certain test items in the different grades. The method previously described assumed that all of the observed proportions of right answers, such as those of Woody, are of equal reliability. As is well known, however, the sigma value of a test question which is answered by ninety-eight per cent of Grade V children, for example, is not nearly so reliable as the sigma value of a test item that is answered correctly by fifty per cent of the children in that grade. The weight assigned to any observation should be inversely proportional to the square of its standard error or its probable error. It is our present object to extend the absolute scaling method previously described so that the sigma values will be given weights in accordance with their reliabilities and so that the most representative scale like that of Fig. 1 may be constructed. With such a scale established it is possible to ascertain, if desired, the weighted average scale value of each test item as determined by the several age or grade groups.

The Obersvation Equation.—Let any two adjacent frequency curves in Fig. 1 represent two overlapping grade groups and let them be designated by the subscripts $_1$ and $_2$. Let M_1 and M_2 be the scale values of the means of the two groups and let σ_1 and σ_2 represent their respective dispersions on the absolute scale. If any particular test item or a particular raw score is allocated on the absolute scale, its scale value should be ideally the same whether determined by group one or by group two. Let x_1 be the sigma value of the proportion, P_1, of group one that exceeds a given score, and let x_2 have a similar meaning for the same test score in group two. Then, by the logic explained more fully in the previous article, we have as the first form of the observation equation

$$M_2 + x_2\sigma_2 = M_1 + x_1\sigma_1 \qquad (1)$$

which may be written

$$x_2 = x_1 \frac{\sigma_1}{\sigma_2} + \frac{(M_1 - M_2)}{\sigma_2} \tag{2}$$

If we plot x_2 against x_1 we should have a linear plot in which

$$\text{the slope} = \frac{\sigma_1}{\sigma_2} \tag{3}$$

and the

$$x_2\text{-intercept} = \frac{(M_1 - M_2)}{\sigma_2} \tag{4}$$

The usual least square methods are not applicable for finding the constants in equation (2) because neither x_2 nor x_1 can be considered as the independent or the dependent variable. By considering x_1 and then x_2 as the independent variable, there will be two least square solutions which correspond to the two regression lines for the linear plot of equation (2). In order to have a unique solution which shall assign errors of measurement to both x_1 and x_2 we adopt the line of relation which is the regression line for the plot with the correlation coefficient taken as unity. Hence the equation of the line of relation will be

$$y = \frac{\sigma_y}{\sigma_x} x$$

which becomes

$$y = \frac{\sigma_y}{\sigma_x} x \tag{5}$$

when $r = 1.00$ and, in terms of the original observed quantities, this equation becomes

$$Y - m_y = \frac{\sigma_y}{\sigma_x}(X - m_x) \tag{6}$$

Now let S_1 = standard deviation of x_1
S_2 = standard deviation of x_2
m_1 = means of x_1
m_2 = mean of x_2.

Rewriting equation (6) with the present notation, we have

$$x_2 - m_2 = \frac{s_2}{s_1}(x_1 - m_1) \tag{7}$$

This can be written

$$x_2 = \frac{s_2}{s_1} x_1 + \left(m_2 - \frac{s_2}{s_1} \cdot m_1\right) \tag{8}$$

This is a more useful form of the observation equation (1).

Plotting x_2 against x_1, we should have a linear relation in which

$$\text{the slope} = \frac{s_2}{s_1} \tag{9}$$

$$x_2\text{-intercept} = \left(m_2 - \frac{s_2}{s_1}m_1\right) \tag{10}$$

But since equations (2) and (8) represent the same plot, it follows that (3) and (9) are identical and that (4) and (10) are identical.

Hence

$$\text{the slope} = \frac{s_2}{s_1} = \frac{\sigma_1}{\sigma_2} \tag{11}$$

or

$$\sigma_2 = \sigma_1\frac{s_1}{s_2} \tag{12}$$

and the

$$x_2\text{-intercept} = \left(m_2 - \frac{s_2}{s_1}m_1\right) = \frac{M_1 - M_2}{\sigma_2} \tag{13}$$

By substitution from equation (11) this can be re-written as

$$M_2 = \sigma_1 m_1 - \sigma_2 m_2 + M_1 \tag{14}$$

Equations (12) and (14) are the fundamental equations for our scaling purposes. The problem consists, then, in finding the values of σ_2 and M_2. These values may be determined from equations (12) and (14) when σ_1 and M_1 have assumed values which determined the arbitrary origin and the unit of measurement for the absolute scale.

In order to ascertain the value of σ_2 by equation (12) we must know the numerical value of the ratio between the two s-values.

The deviation of x_1 from its mean m_1 is

$$(x_1 - m_1)$$

The weighted squared deviation is therefore

$$w_{12}(x_1 - m_1)^2$$

The determination of the weight w_{12} will be shown in a later section of this paper.

The standard weighted deviation of x_1 is therefore

$$s_1 = \sqrt{\frac{\Sigma[w_{12}(x_1 - m_1)^2]}{\Sigma w_{12}}} \tag{15}$$

In order to facilitate the calculation of s_1 we make the following transformations.

$$\begin{aligned}\Sigma[w_{12}(x_1 - m_1)^2] &= \Sigma[w_{12}(x_1^2 - 2m_1x_1 + m_1^2)] \\ &= \Sigma[w_{12}x_1^2 - 2w_{12}m_1x_1 + w_{12}m_1^2] \\ &= \Sigma(w_{12}x_1^2) - 2m_1\Sigma w_{12}x_1 + m_1^2\Sigma w_{12}\end{aligned}$$

$$\frac{\Sigma[w_{12}(x_1 - m_1)^2]}{\Sigma w_{12}} = \frac{\Sigma(w_{12}x_1^2) - 2m_1\Sigma w_{12}x_1 + m_1^2\Sigma w_{12}}{\Sigma w_{12}}$$

But

$$m_1 = \frac{\Sigma w_{12}x_1}{\Sigma w_{12}}$$

Hence

$$\frac{\Sigma[w_{12}(x_1 - m_1)^2]}{\Sigma w_{12}} = \frac{1}{(\Sigma w_{12})^2}[\Sigma w_{12}x_1^2 \cdot \Sigma w_{12} - (\Sigma w_{12}x_1)^2]$$

$$\therefore s_1 = \frac{1}{\Sigma w_{12}}\sqrt{\Sigma w_{12}x_1^2 \cdot \Sigma w_{12} - (\Sigma w_{12}x_1)^2} \qquad (16)$$

By analogy

$$s_2 = \frac{1}{\Sigma w_{12}}\sqrt{\Sigma w_{12}x_2^2 \cdot \Sigma w_{12} - (\Sigma w_{12}x_2)^2} \qquad (17)$$

and therefore

$$\frac{s_2}{s_1} = \sqrt{\frac{\Sigma w_{12} \cdot \Sigma w_{12}x_2^2 - (\Sigma w_{12}x_2)^2}{\Sigma w_{12} \cdot \Sigma w_{12}x_1^2 - (\Sigma w_{12}x_1)^2}} \qquad (18)$$

This form enables us to calculate the numerical value of s_2/s_1.

The numerical value of M_2 can be ascertained by equation (14) after the two means m_1 and m_2, are known. The weighted means of x_1 and x_2 are calculated by the following relations:

$$m_1 = \frac{\Sigma w_{12}x_1}{\Sigma w_{12}} \qquad (19)$$

$$m_2 = \frac{\Sigma w_{12}x_2}{\Sigma w_{12}} \qquad (20)$$

To ascertain the Weight w_{12} of the Observation Equation.—The observation equation (1) for absolute scaling can be written

$$Q = (M_1 - M_2) - x_2\sigma_2 + x_1\sigma_1 \qquad (21)$$

in which Q is a residual.

Let E = standard error of Q

e_1 = standard error of x_1

e_2 = standard error of x_2

W = weight of the observation equation (21).

Then[1]

$$E = \sqrt{\left(\frac{\partial Q}{\partial x_1}\right)^2 e_1^2 + \left(\frac{\partial Q}{\partial x_2}\right)^2 \cdot e_2^2} \tag{22}$$

But

$$\left(\frac{\partial Q}{\partial x_1}\right)^2 = +\sigma_1^2 \tag{23}$$

and

$$\left(\frac{\partial Q}{\partial x_2}\right)^2 = +\sigma_2^2$$

$$e_1 = \frac{\sigma_1}{Z_1}\sqrt{\frac{p_1 g_1}{n_1}} \tag{24}$$

in which p_1 = the observed proportion of correct answers to any specified test item.

$q_1 = 1 - p_1$

n_1 = the number of cases on which p_1 is determined.

σ_1 = the standard deviation of the distribution of test ability in the given age group.

z_1 = the ordinate of the distribution which divides the surface into the two parts, p_1 and q_1, when σ_1 and n_1 are both taken as unity. It is the entry Z in the Kelley-Wood tables.[2]

$$\therefore e_1^2 = \frac{\sigma_1^2 p_1 q_1}{z_1^2 n_1} \tag{25}$$

and, by analogy,

$$e_2^2 = \frac{\sigma_2^2 p_2 q_2}{z_2^2 n_2}$$

Substituting from (23) and (25) in (22) we have

$$E = \sqrt{\frac{\sigma_1^4 p_1 q_1}{z_1^2 n_1} + \frac{\sigma_2^4 p_2 q_2}{z_2^2 n_2}} \tag{26}$$

The weight W of the observation equation (1) is

$$w = \frac{1}{E^2} \tag{27}$$

Therefore

$$W = \frac{1}{\frac{\sigma_1^4 p_1 q_1}{z_1^2 n_1} + \frac{\sigma_2^4 p_2 q_2}{z_2^2 n_2}} \tag{28}$$

---

[1] See Weld, L. D.: "Theory of Errors and Least Squares." P. 165, equation 102.

[2] Kelley, T. L.: "Statistical Method." P. 90, equation 43.

Since σ_1 and σ_2 are at least roughly comparable for adjacent age groups and since n_1 and n_2 are usually comparable for adjacent age groups, these four factors may be ignored for purposes of weighting. With this simplification

$$W = \frac{1}{\frac{p_1q_1}{z_1^2} + \frac{p_2q_2}{z_2^2}} \tag{29}$$

To facilitate calculation, the values

$$v_1 = \frac{p_1q_1}{z_1^2} \tag{30}$$

and

$$v_2 = \frac{p_2q_2}{z_2^2}$$

may be determined once for all by Kelley's tables and arranged in tabular form, showing the numerical value of v for each given value of p. This can readily be done because it is seen from (30) that v_1 is completely determined by p_1.

For purposes of calculation

$$W = \frac{1}{v_1 + v_2} \tag{31}$$

for which the values of v_1 and v_2 are read off from a facilitating table. Table I has been prepared to show the value of v for each value of p

TABLE I.[1]—TABLE GIVING VALUES OF $\frac{pq}{z^2}$ FOR VALUES OF p

	0	10	20	30	40	50	60	70	80	90
0	∞	2.92	2.04	1.74	1.61	1.57	1.61	1.74	2.04	2.92
1	14.14	2.77	2.00	1.72	1.60	1.57	1.62	1.76	2.09	3.10
2	8.52	2.64	1.96	1.70	1.59	1.57	1.62	1.78	2.14	3.33
3	6.33	2.53	1.92	1.69	1.59	1.57	1.63	1.80	2.20	3.62
4	5.19	2.43	1.89	1.67	1.58	1.58	1.65	1.83	2.27	3.97
5	4.48	2.34	1.86	1.66	1.58	1.58	1.66	1.86	2.34	4.48
6	3.97	2.27	1.83	1.65	1.58	1.58	1.67	1.89	2.43	5.19
7	3.62	2.20	1.80	1.63	1.57	1.59	1.69	1.92	2.53	6.33
8	3.33	2.14	1.78	1.62	1.57	1.59	1.70	1.96	2.64	8.52
9	3.10	2.09	1.76	1.62	1.57	1.60	1.72	2.00	2.77	14.14

[1] These values were calculated with the aid of the Kelley-Wood Table. Example: When p is .72, the value of v is 1.78.

to the second decimal. The values have been calculated by means of the Kelley-Wood tables. A similar table can of course be prepared for values of p to three decimals but such refinement is not necessary merely for purposes of weighting, especially in view of the fact that equation (29) already contains two approximations. The errors of Table I become largest for the extreme values of p and it is these extreme values of p that have low weights.

Application of the Method to the Woody Arithmetic Scales.—In order to construct Fig. 1 which represents Woody's arithmetic data the distribution for the second grade was drawn first. Its mean is designated zero and its standard deviation unity. The problem is, then, to draw the distribution for the third grade so that it will be consistent with that of the second grade. This necessitates the numerical determination of M_3 and σ_3. When these have been determined, the same process is repeated for the determination of M_4 and σ_4 with M_3 and σ_3 as known, and so on for the successive grade groups.

Table II shows a convenient arrangement of the data. The subscript $_1$ is used for the lower group in each grade combination, in this case Grade III. The subscript $_2$ is used for the upper group in each grade combination. Here it represents Grade IV. The first column shows merely the code numbers of Woody's arithmetic problems. Columns P_1 and P_2 are copied directly from Woody's Table XV, page 52. The columns x_1 and x_2 are the sigma values corresponding to P_1 and P_2. The entries v_1 and v_2 are obtained from Table I. Thus P_1 for problem No. 1 is .742 and in Table I it will be found that v for .74 is 1.83. The values of v are given in Table I only for two decimals of p. The entry v_2 is found in the same way. The entry $(v_1 + v_2)$ is self-explanatory. The entry w is found by equation (31). The remaining columns should be clear from their captions. It will be noticed that several of the test problems were omitted since they did not overlap Grades III and IV. Finally the five indicated summations are recorded on this data sheet. These summations constitute the objective in Table II.

The calculations are next arranged in an outline as shown in Table III. The outline facilitates accuracy and checking. It is divided into six sections. In section (1) are recorded the two adjacent grade groups, III and IV, and the subscripts of the general formulæ to which they apply. Here are also recorded the values of σ_1 and M_1 for the lower grade group which is in this case Grade III.

TABLE II

No.	p_1	p_2	x_1	x_2	v_1	v_2	$v_1 + v_2$	wx	wx_1	wx_2	wx_1^2	wx_2^2
1	.742	.829	− .65	− .65	1.83	2.20	4.03	.2481	−.16127	−.23570	.10482	.22391
2	.946	.989	+1.61	−2.29	4.48	14.14	18.62	.0537	−.08646	−.12297	.13920	.28161
3	.964	.989	−1.80	−2.29	5.19	14.14	19.33	.0517	−.09306	−.11839	.16751	.27112
4	.761	.866	− .71	−1.11	1.89	2.53	4.42	.2262	−.16060	−.25108	.11403	.27870
5	.815	.921	− .90	−1.41	2.14	3.33	5.47	.1828	−.16452	−.25775	.14807	.36342
6	.862	.939	−1.09	−1.55	2.43	3.97	6.40	.1563	−.17037	−.24227	.18570	.37551
7	.875	.939	−1.15	−1.55	2.64	3.97	6.61	.1513	−.17400	−.23452	.20009	.36350
8	.771	.901	− .74	−1.29	1.92	2.92	4.84	.2066	−.15288	−.26651	.11313	.34380
9	.862	.944	−1.09	−1.59	2.43	3.97	6.40	.1563	−.17037	−.24852	.18570	.39514
10	.864	.947	−1.10	−1.62	2.43	4.48	6.91	.1447	−.15917	−.23441	.17509	.37975
11	.883	.959	−1.19	−1.74	2.64	5.19	7.83	.1277	−.15196	−.22220	.18084	.38662
12	.649	.899	− .38	−1.28	1.66	2.92	4.58	.2183	−.08295	−.27942	.03152	.35766
13	.373	.620	+ .32	− .31	1.63	1.62	3.25	.3077	+.09846	−.09539	.03151	.02957
14	.659	.899	− .41	−1.28	1.67	2.92	4.59	.2179	−.08934	−.27891	.03663	.35701
15	.533	.829	− .08	− .95	1.57	2.20	3.77	.2653	−.02122	−.25204	.00170	.23943
16	.387	.643	+ .29	− .37	1.62	1.65	3.27	.3058	+.08868	−.11315	.02572	.04186
17	.468	.716	+ .08	− .57	1.57	1.78	3.35	.2985	+.02388	−.17015	.00191	.09698
18	.338	.642	+ .42	− .36	1.67	1.65	3.32	.3012	+.12650	−.10843	.05313	.03904
19	.150	.409	+1.04	+ .23	2.34	1.60	3.94	.2538	+.26395	+.05837	.27451	.01343
20	.142	.510	+1.07	− .03	2.43	1.57	4.00	.2500	+.26750	−.00750	.28623	.00023
21	.116	.459	+1.20	+ .10	2.64	1.58	4.22	.2370	+.28440	+.02370	.34128	.00237
22	.080	.339	+1.41	+ .42	3.33	1.67	5.00	.2000	+.28200	+.08400	.39762	.03528
23	.006	.056	+2.51	+1.59	14.14	3.97	18.11	.0552	+.13855	+.08777	.34777	.13955
25	.005	.023	+2.58	+2.00	14.14	8.52	22.66	.0441	+.11378	+.08820	.29355	.17640
31	.005	.056	+2.58	+1.59	14.14	3.97	18.11	.0552	+.14242	+.08777	.36743	.13955
34	.024	.213	+1.98	+ .80	8.52	2.00	10.52	.0951	+.18830	+.07608	.37283	.06086
35	.006	.095	+2.51	+1.31	14.14	2.92	17.06	.0586	+.14709	+.07677	.36919	.10056

$\Sigma w = 4.8691.$
$\Sigma wx_1 = .32734.$
$\Sigma wx_2 = -3.15665.$
$\Sigma wx_1^2 = 4.94671.$
$\Sigma wx_2^2 = 5.49286.$

TABLE III

(1) Subscript 1 refers to Grade III
Subscript 2 refers to Grade IV

$\sigma_1 = 1.69758$
$M_1 = 1.59239$

(2)

$\Sigma w = 4.8691$
$\Sigma wx_1 = .32734$
$\Sigma wx_2 = -3.15665$
$\Sigma wx_1^2 = 4.94671$
$\Sigma wx_2^2 = 5.49286$
$(\Sigma wx_1)^2 = .10715$
$(\Sigma wx_2)^2 = 9.96444$

$$m_1 = \frac{\Sigma wx_1}{\Sigma w} = .06723$$

$$m_2 = \frac{\Sigma wx_2}{\Sigma w} = -.64830$$

(3) $$\frac{S_1}{S_2} = \sqrt{\frac{\Sigma w \cdot \Sigma wx_1^2 - (\Sigma wx_1)^2}{\Sigma w \cdot \Sigma wx_2^2 - (\Sigma wx_2)^2}} = \sqrt{\frac{(4.8691)(4.94671) - .10715}{(4.8691)(5.49286) - 9.96444}}$$
$$= 1.19538$$

(4) $$\frac{S_2}{S_1} = .83655$$

$$\sigma_2 = \sigma_1 \cdot \frac{S_1}{S_2} = (1.69758)(1.19538)$$
$$= 2.02925$$

(5) $$M_2 = \sigma_1 m_1 - \sigma_2 \cdot m_2 + M_1$$
$$= 3.02207$$

(6) $$x_2 = x_1 \cdot \frac{s_2}{s_1} + \left(m_2 - m_1 \cdot \frac{s_2}{s_1}\right)$$
$$x_2 = .83655x_1 - .70454$$

TABLE IV.—VALUES OF M AND σ FROM WOODY DATA

GRADE	M	σ
II	0.00000	1.00000
III	1.59239	1.69758
IV	3.02207	2.02925
V	4.81116	2.58900
VI	7.09823	3.53254
VII	8.61599	4.15437
VIII	10.09108	4.52157

In section (2) are copied the five summations from Table II. Four additional items are here recorded which are derived from the five summations.

In section (3) we calculate the numerical value of the ratio of the two dispersions by equation (18).

In section (4) are recorded the reciprocal of this ratio and equation (12) which enables us to calculate the dispersion of Grade IV.

In section (5) we have equation (14) which enables us to calculate the mean scale value for Grade IV.

Section (6) contains equation (8) of the linear plot of Fig. 2. This constitutes a graphic check on our work. If equation (8) when plotted directly on Fig. 2 seems to fit the observations it is a reasonable guarantee that at least no gross errors have been made.

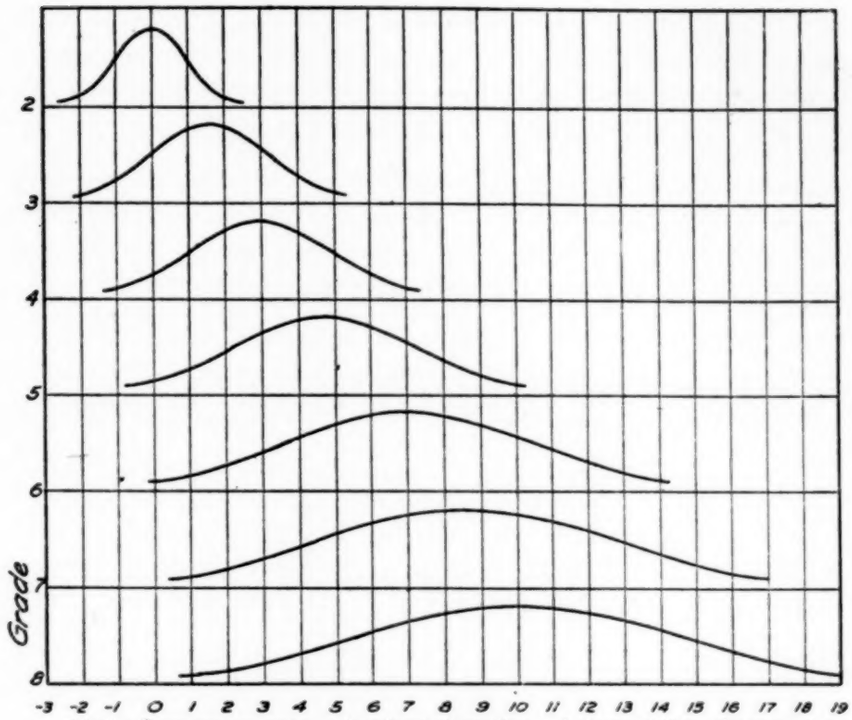
Grade
2
3
4
5
6
7
8
-3 -2 -1 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
Unit of measurement: standard deviation of test ability of second grade children
Origin of measurement: mean test ability of second grade

FIG. 1.

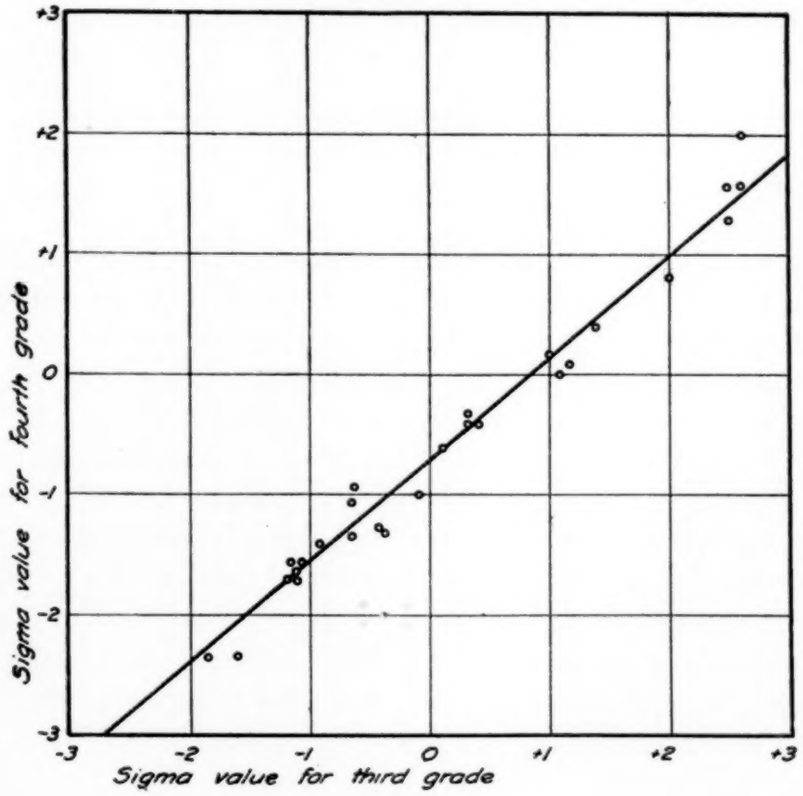
Sigma value for fourth grade
+3
+2
+1
0
-1
-2
-3
-3 -2 -1 0 +1 +2 +3
Sigma value for third grade

FIG. 2.

Since there are seven grade groups in Woody's data, it was necessary to carry out six sets of calculations like those here shown as samples, namely the grade combinations II-III, III-IV, IV-V, V-VI, VI-VII, VII-VIII. The calculations for the grade combination III-IV is here used as an example. When these six sets of calculations have been completed we can prepare Table IV which shows the mean scale value and the dispersion for each of the seven grades in Woody's data. Figure 1 is then drawn by means of the data of Table IV.

An interesting conclusion from the present analysis of Woody's data is that absolute variability increases markedly with grade. Figure 3 shows absolute variability in the Woody arithmetic tests

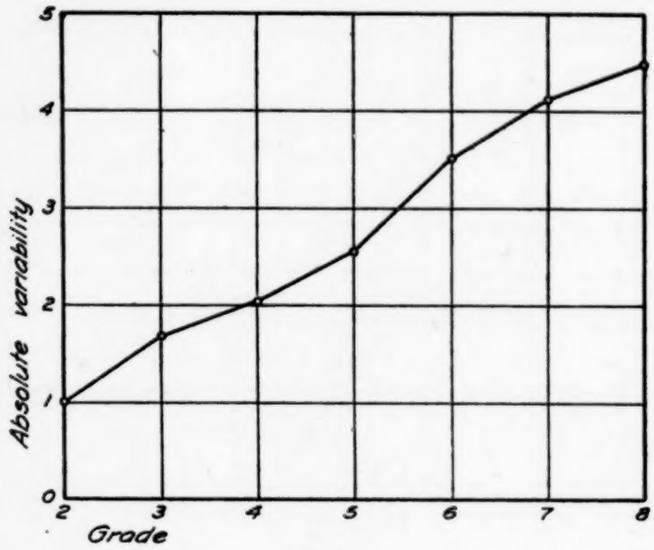

FIG. 3.

plotted against grade. It appears that the variability in Grade VIII is 4.52 times that of Grade II. This striking increase in variability of arithmetical ability through the grades appears also in Fig. 1 in the rapid increase in the dispersions of the successive grades.

Summary.—A set of fundamental scaling formulæ have been developed by which absolute scaling can be accomplished with weighted observations. The original observation equation (1) is rewritten in more useful form for purposes of calculation as equation (8). When the standard deviation of one of the age or grade groups is known or assumed, the standard deviation of the other may be calculated by means of equation (12). When the mean test ability of one of the age groups is known or assumed, the mean test ability of the other can be calculated by equation (14). The ratio of the weighted

dispersions is needed in equation (12) and it can be calculated in terms of the original data by equation (18). The approximate weight for each observation equation can be calculated by equation (29) or, better, with the aid of a facilitating table such as Table I by equation (31).

The detailed outline of calculation is shown in Tables II and III for the grade combination III-IV of Woody's data on arithmetic tests. Similar computations were made for the six other grade combinations of Woody's data. The result is Table IV which shows the mean and the standard deviation of each grade in his arithmetic tests on the absolute scale. These results are reproduced graphically in Figs. 1 and 3.

One conclusion of some interest is the striking increase in absolute variability through the grades which is immediately apparent in Figs. 1 and 3. It should be noted that this finding contradicts the assumptions in Woody's scaling method by which the dispersions are assumed to be constant in all the grades. This false assumption has been previously pointed out in connection with Trabue's Language Completion scales which are also invalidated by the same assumption.[1] It is not here made a major issue because of its previous description in connection with Trabue's language scales. Our principal objective here has been to show how absolute scaling may be accomplished with weighted observations.

[1] Thurstone, L. L.: The Unit of Measurement in Educational Scales. To be published in *The Journal of Educational Psychology.*

LECTURE AND CLASS DISCUSSION IN TEACHING EDUCATIONAL PSYCHOLOGY

RALPH B. SPENCE

Teachers College, Columbia University, N. Y. C.

In considering the problem of method, or as I prefer to phrase it, the problem of the most effective kind of working relationship between student, instructor, and subject, certain checks must be made upon our lines of reasoning. It is one thing to discover the generalizations about individual differences, learning, responses to other individuals, and the like, and it is another to apply these generalizations to the concrete situation at hand. One can say, on the basis of his knowledge of psychology, that this method would seem to be desirable and that method would seem wasteful. He must go further however and find out to what extent this reasoning will still hold true in college *A* with an endowment of *B* dollars; classroom facilities *C;* students *D*, *E*, *F*, . . . whose previous experiences have been *G*, *H*, *I*; professors *J*, *K* . . . each having spent *L*, *M*, years in presenting his subject *N*, *O* in a particular way. The first method of attack shows us the direction in which we must strive to move, the second indicates the present first move. One may, in starting for a point due north, get there most effectively by moving first toward the east or even the southeast.

With these considerations in mind, we can be surer that the results which follow will be interpreted in light of the specific situation from which they came. The present experiment is one of the attempts which Dr. Goodwin B. Watson and the author have made to secure better results from the teaching of Educational Psychology. The lecture method did not seem to recommend itself as wholly desirable and yet, as it is such a common type of teaching it certainly was desirable to compare other plans with it. The experiment, therefore, set out to compare the lecture method with a plan involving class discussion.

Studies in connection with the problem of teaching methods are not numerous. Jones[1] in an experimental study of college teaching found that the lecture method was not effective from the standpoint of permanence of learning. Scheidemann[2] has reported that the lecture-conference method and individualized instruction seemed equally

[1] *Archives of Psychology*, 1925.

[2] *School and Society*, Vol. XXV, 1927, pp. 672–674.

effective for elementary psychology, but English[1] has called attention to several factors which must be kept in mind when the results are interpreted.

The experimental group in this study consisted of a class in Educational Psychology, meeting two hours on Saturday mornings, and continuing over both semesters. Students were permitted, however, to take either semester without the other and there was a large change in February, many students dropping out and others entering. The students were, in part, full time graduate students and, in part, students taking only a few courses. Many of the latter were teachers in and around New York City.

No announcement was made in the catalog about sections, and on the first day everyone registered for the course came to one room and was then assigned to one of the two sections. This made it possible to section the class in any way desired. Each section was large, the enrollment being about as follows:

	Semester I	Semester II		
		Old	New	Total
Section A	172	66	59	125
Section B	153	62	89	151

The experimental plan was as follows:

Section A— Initial Test—Lecture-test$_2$—Discussion-test$_3$.

Section B—Initial Test—Discussion-test$_2$—Lecture-test$_3$.

The same instructor remained with the section throughout the year, lecturing during one semester and discussing during the other.

The initial test and the test$_3$ were the same except that each student took a different form. In September, one-half of each class was given Form C and the other half Form D, and in May the forms were reversed. The test was designed to include application as well as knowledge by means of such questions as:

A four year old child has bad spells of temper tantrums. Classify each suggested methody of treatment as (1) desirable, (2) likely to make little difference one way or another, (3) undesirable.

(*a*) Grow angry at the child. ( )

(*b*) Ignore the child. ( )

[1] *School and Society*, Vol. XXVI, 1927, pp. 172–174.

(*c*) Give the child what he wants to stop the fuss. ( )
(*d*) See that the child does not get what it wants through tantrums. ( )
(*e*) Give the child a whipping whenever it starts a tantrum. ( )

How significant are each of the following elements in providing for efficient study? Answer: (1) Important, (2) desirable but less important (3) sometimes desirable, sometimes undesirable, (4) undesirable.

(*a*) A strong interest or motive. ( )
(*b*) Outlining or summarizing material read. ( )
(*c*) Absence of strong emotion. ( )
(*d*) Moderate temperature, clear dry atmosphere. ( )
(*e*) Practice in recitation or other use of material studied. ( )
(*f*) Control of daydreams. ( )

We have no knowledge as to just what such questions test and how they would compare with questions which were phrased without application to actual problems. The test was limited to forty-five minutes and almost no one finished, so that factors such as speed of reading in part influenced the results.

The test for the end of the first semester was composed of two parts and the total time allowed was seventy minutes, so although the total test was longer than the other, speed probably played a less important part. This would seem to be verified by the correlation between the scores on this test and the scores on Otis Self-administering Intelligence Test, shown later.

Description of the Methods

In the lecture method, the instructor talked for two fifty minute periods, allowing time at the end of the period for answering questions. These questions were either handed in beforehand or asked directly at that time. The average time for questions, however, was only about five minutes per week. A very comprehensive bibliography (similar to the one included in "Sketches In and Out of School") was mimeographed and given to each student. During the first semester the students handed in reports of their readings on three by five cards but no definite amount was required. In the second semester, a regular reading report form was prepared and students were required to hand in at least one a week. This increased the amount of reading considerably. Four shorter quizzes were given during each semester and reports on cases and abstracts of magazine articles were also included.

In the discussion method the general plan was to take a case similar to those published in "Sketches In and Out of School" illustrating the particular topic which was being studied and center the discussion

around the case. For example, when intelligence tests were being studied, intelligence test data for an elementary school class were presented. The case was analyzed to see what information was necessary, various class members gave reports of their readings and practical experiences, and then the group tried to state their conclusions, which were summarized by the instructor at the end of the period. Each student was required to hand in each week a short paper on some topic which would help him begin thinking in the direction of the discussion for that week. In the case of the above example, the class was asked to put down all the things which tended to determine the score an individual made in an intelligence test and to evaluate the importance of each. The quizzes and reading reports were the same for the discussion group as for the lecture group.

The Results

Because of the large change in the personnel of the class at the end of the first semester, it was necessary to confine the study to individuals who were included in both groups. When these groups were reduced to the students for whom complete records were available, there were sixty students in each section. Our results must therefore be reported in terms of sixty students, but it must be kept in mind that these sixty were actually only a part of the group.

The results for the three tests and for the Otis Self-administering were T-scaled separately.[1] This meant that the average for the one hundred twenty students on any one of the four tests would be fifty and the average for each section would enable comparisons to be made easily.

	Initial test	Middle test	Final test	Otis S.-A.
Section A				
Mean	48.9	53.4	50.1	50.1
SD	9.78	8.85	8.49	10.15
SD of mean	1.26	1.14	1.10	1.35
Section B				
Mean	51.1	46.2	49.9	49.6
SD	9.18	8.55	10.38	9.18
SD of mean	1.19	1.10	1.38	1.19

[1] McCall: "How to Measure in Education."

	Differences
1st Semester.	
Section A (lecture) gained	9.4 units over Section B (discussion)
SD difference =	1.58
2nd Semester.	
Section B (lecture) gained	7.0 units over Section A (discussion)
SD difference =	1.76

The results of the tests show therefore that these students tended to do better work in lectures at least as far as measurable results on examinations is concerned. During the first semester, the students

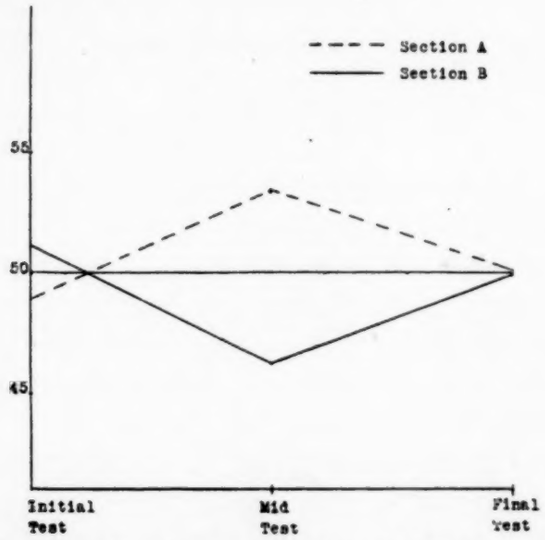

Fig. 1.—Relative changes of sections (in T-score values).

who were having lectures forged ahead of those who were having discussions and during the second semester the students in the lecture group regained the lost ground. Statistically the differences are significant. In practical terms, the superiority of the lecture group corresponds to the difference between a grade of B and a grade of C, for example. If we assume a distribution of marks as follows: A, ten per cent; B, twenty per cent; C, forty per cent; D, twenty per cent; F, ten per cent; and compute T-scores for each letter, we get the following values: A, sixty-six; B, fifty-eight; C, fifty; D, forty-two; F, thirty-four.

This difference between letter grades of eight points is about the average difference between lecture and discussion methods. The correlation between the various tests are as follows:

	I	*M*	*F*	*O*
Initial test—Section A	...	.28	.41	.44
B	...	.51	.52	.57
Middle test—Section A	...	...	.44	.33
B	...	...	.20	.35
Final test—Section A	...	...	...	.43
B	...	...	...	.64
Otis S.-A.—Section A				
B				

PE of an *r* of .60 for this group would be ± .06
PE of an *r* of .50 for this group would be ± .07
PE of an *r* of .40 for this group would be ± .07
PE of an *r* of .30 for this group would be ± .08
PE of an *r* of .20 for this group would be ± .08

Discussion

There are a large number of questions which must be raised. Some of the points to be considered are:

(*a*) Aims desired.
(*b*) Failure to measure significant changes.
(*c*) Size of group.
(*d*) Previous and parallel training of group.
(*e*) Mechanical arrangements—time of meeting, reading facilities, etc.
(*f*) Interest and ability of instructor.
(*g*) Other possible alternatives.

The verdict of the effectiveness of any place for teacher-pupil relationship depends upon the weight one assigns to the various changes desired. I do not think that one would find much difference between objectives listed by various instructors but one would find real differences between the weights allotted to these objectives. One person is more anxious that his group know certain principles of learning than he is that the individuals improve their ability to work together; another holds that an attitude of open-mindedness is far more valuable than ability to pass a test on a textbook.

The problem of aims is closely connected with the ability to measure. Unless one has a fairly reliable test to indicate how much he is accomplishing along various lines, it would be wisest to give most weight to the things which can be measured reliably. In the particular

instance it is probably quite true that a more significant measure of the differences between the two plans would come from practical tests given a year or two after the course had been completed. Under present conditions, however, we have no way of doing that.

In connection with attitudes, we have some information from student estimates. Each student was asked to write out his reactions to the course and this included for those students who took both semesters a comparison of the two plans. It proved however that the students were too close to the experiment to be able to pass judgment. In the section which had the lecture method last, the preference was two to one in favor of the lecture. In the other section the preference was about three to two in favor of discussion. There is no definite indication from the comments that students felt the development of open-mindedness was greater under one plan than under another.

As was pointed out above, the classes were in all cases large. The smallest class had more than one hundred twenty; the largest had one hundred seventy. It probably is true that discussion in large groups is much more difficult than in small groups and much less satisfying. The problem is complicated by the attitudes of the group. If it were possible to have a group strongly desiring the discussion method, mere size might not be the limiting factor. More information is needed on the effect of class size.

The previous training of the students was almost devoid of useful experience in group discussion. Some of them were getting somewhat similar experiences in other classes, but it was incidental and undirected training. The result was that many students had a distinct mind-set for lectures and a marked antagonistic feeling towards other students who used class time to relate their experiences. It is quite true that in a number of instances the discussion dragged, it was impossible to hear, and students could not be kept directly to the point.

Such conditions are on the whole, however, not unlike those met with in many classes, but the solidifying effect of having previously been trained almostly solely by the lecture method is probably greater in this group. Many of the students during the week teach by authority and themselves honor the method on Saturday. The teacher speaks! Everybody must listen carefully.

The mechanical arrangements have been already partly discussed, but other points may be added. This particular group met only once a week and many of the students were not in a position to do much outside reading. In a similar course which meets three times

a week, the feeling toward the discussion plan is better. Proper conduct of a discussion especially for a beginning group means ready access to source material and ability to get guidance from the instructor. Rooms arranged as all classrooms are at present are also not conducive to good group feeling. A situation where students knew one another well would produce different results from those where one student knows almost nothing about ninety-five per cent of the class members.

As far as the instructors involved in this experiment are concerned, on the whole they tend to prefer the discussion method to lecturing. Furthermore Dr. Watson has had considerable experience with discussion groups so that the criticisms which might be made that discussion group method was inadequately presented has less weight than it would if there had been no such background of experience for either instructor. As far as I know, there is no information which would help us decide the effect of the instructor on the results or to what extent one can expect the average individual to improve his ability at leading the discussion by proper training.

The final consideration is that perhaps after all, the way to handle the situation is by some combination of methods. For experimental reasons we avoid use of lecture method during the discussion phase. There are many other possible ways of handling discussion groups which might be tried out. One might also try making a very definite set of assigned readings and mimeographed material (similar to lecture material) to be given to each pupil at the beginning of the year. Each student would then only come to class for quizzes. All of this is useful here only to show the limits of the present experiment.

Conclusion

In the light of the above discussion, we can conclude that for two large (N = 120–170) sections of graduate students in Educational Psychology, meeting once a week for two hours, with little previous experience or feeling in favor of the discussion method, the lecture method is superior to the discussion method in producing improvement in things measured by the tests. The major value of the experiment is to show the high degree of specificity of conditions in connection with college teaching and to point out the lack of information at present available. Some of the factors which must be considered are:

(*a*) Weight given to training along various lines—knowledge, attitudes, ability to work with one another.

(*b*) Size of group.

(*c*) Previous and parallel training of the group.
(*d*) Mechanical conditions which must be met.
(*e*) Variations of method.
(*f*) Instructor—ability, interest, training.

The problem that must be solved is not the question, "Is Method *A* better than Method *B* or Method *C*?" but rather, "What are the conditions under which Method *A* produces most effective results? What are the situations where Method *B* is best? etc." There are times when the lecture method is an effective aid in securing desirable changes in students. There are other times when the lecture method is harmful and should be replaced by discussion or individual projects or something else.

The decrying of the wholesale use of lectures in college teaching is probably justified. The wholesale decrying of the use of lectures in college teaching is just as certainly not justified.

THE RETIREMENT OF PROFESSOR EDWIN A. KIRKPATRICK

Professor Edwin A. Kirkpatrick, after thirty years of service in the State Normal School at Fitchburg, Massachusetts, retired in June 1928.

The editors take pleasure in recognizing at this time the important contributions which Professor Kirkpatrick has made in applying modern psychology to educational problems.

Since 1895 when his "Inductive Psychology" was published, he has been a prolific writer. Especially noteworthy are his "Fundamentals of Child Study" (1903), "Genetic Psychology" (1909), "The Individual in the Making" (1911). In addition to these he has published "The Use of Money" (1915), "Fundamentals of Sociology" (1916), "Studies in Psychology" (1918), "Imagination and Its Place in Education" (1919).

Professor Kirkpatrick will continue his residence in Leominster, Massachusetts. One of his chief interests is ethics as applied to sociology and other sciences. During the present year he will be engaged in revising his well known work, "Fundamentals of Child Study." A recent letter tells us that his time will also be occupied with writing in the field of ethics as applied to sociology, with some part-time lecturing in colleges and universities, and with working in his garden.

PRESENT STATUS OF THE QUESTION OF RACIAL DIFFERENCES

DALE YODER

State University of Iowa

The question as to whether one race is inferior to another continues a matter of widespread interest and contention. In spite of the fact that several writers have delivered themselves of finalities which they consider conclusive, there continues to be a lively discussion and a continual reopening of the question to admit additional evidence. Within the past two years there have been a half dozen books presented to the public which deal directly with the question of racial differences, and there have been no less than half a hundred articles portraying the results of twenty or thirty investigations.

There are three distinct viewpoints represented in the current literature. The first accepts the fact of race superiority and inferiority and is interested in re-stating it and usually adducing additional evidence to support the thesis. The second viewpoint considers race inferiority possible but not adequately demonstrated, and is usually concerned in balancing arguments for and against the idea. The third is a skeptical group, highly critical of the means used to demonstrate race inferiority and of the results so obtained and generally insisting upon racial equality.

As a whole, the attitudes of the various writers presenting these three viewpoints of the problem are quite changed from those which the exponents of the same contentions showed ten or even five years ago. The race superiority enthusiasts appear to be on the defensive, speaking generally, and they seem to be striving furiously to amass more of the usual evidence to support their thesis, at the same time ignoring the fact that it is the quality rather than the quantity of evidence which is under fire. Those who hold that racial inferiority is a possibility but an undemonstrated one show the reserve and objectiveness of a group which feels that it has the majority of scientific opinion supporting it. The literature of criticism has the fire of enthusiastic discovery about it, combined with an idealistic zeal aiming at the unseating of what it considers scientific inaccuracy and fallacy.

A great deal of the recent writing represents nothing more than dogmatic assertions of one position or another, apparently intended by the writers to settle this disturbing question once and for all. It is

obvious that such statements are of little interest. There are, however, several new writers and a tremendous mass of additional data with which the student who wishes to keep abreast of the discussion must acquaint himself. It is the purpose of this paper to summarize briefly these recent contributions.

The chief exponents of the theory popularized by Lothrop Stoddard and Madison Grant include Peterson, Mercer, Strachan, Koch and Simmons, Garth and Whatley, Luckey, Seago, Lacy, Davidson, Patrick and Hirsch. It should be said that such a summary grouping is apt to be unfair. There are various degrees to which the men mentioned subscribe to the theory of inherent racial superiority. Some of them are convinced and complacent about it. Others are frankly skeptical of their own findings. There are numerous positions between these two. In general, however, the statements of these writers bear out the central thesis of racial difference.

Hirsch's study[1] is sponsored by the National Research Council, recommended to the Genetic Psychological Monograph editors by Professor William McDougall, and transmitted to them by Professor Carl Murchison. The project, conducted under the direction of Professor McDougall, accepts the findings of Professor Ferguson with regard to the inferiority of negro children and is an attempt to continue the investigations of Professor Brigham to show that the differences discovered are natio-racial or sub-racial. There were five thousand five hundred four testees, Grades I to IX inclusive, all foreign born except the Americans and Negroes, all took the tests voluntarily, and principal reliance was placed on the Pintner-Cunningham Primary Mental Test. Their findings ranked the participants as follows as to mentality as measured by the tests: Polish Jews, Swedes, English, Russian Jews, Germans, Americans, Lithuanians, Irish, British Canadians, Russians, Poles, Greeks, Italians, French Canadians, Negroes, Portuguese. The Negro IQ, it is interesting to note in connection with other studies to be mentioned shortly, was found to be 84.6. They conclude that the language factor is negligible,[2] that there are grades of intelligence among the children of various occupational groups so that "intelligence is related to occupation as cause is to effect,"[3] and

---

[1] Hirsch, N. D. M.: "Genetic Psychological Monographs." Vol. I, 1926, pp. 239–406.

[2] *Ibid.*, p. 313.

[3] *Ibid.*, p. 328.

that "the correlation of anthropometric characteristics and intelligence is discovered to approach zero."[1] Their conclusion with regard to the race problem in the United States is summarized in the following paragraph:

> Continued segregation or miscegenation—neither of these alternatives should be tolerated by the intelligent and energetic Americans of today. Another alternative, the formation of a Negro State fostered and protected by the United States and perhaps Great Britain, is the only solution of this problem, even though to many, at first thought, it may seem impossibly fantistic.[2]

James Ruey Patrick, Phelps-Stokes fellow in the University of Georgia, has made a study[3] of forty-seven Negroes and forty-seven whites in the public schools of Athens, Georgia. His findings, using the Otis Self-administering Intermediate Examination, Form A, and another examination of his construction for the measurement of ideals, may be thus cataloged:

1. The mean age of the Negro was found to be one year and six months greater than the mean age of the whites.
2. The mean intelligence score of the Negro is twenty, while that of the whites is forty-two.
3. The median IQ of the Negroes is seventy-two; that of the whites is ninety-eight.
4. The whites are only slightly superior in their ideals.
5. The whites are superior in achievement.

L. D. Lacy, Director of Research in the public schools of Oklahoma City, Oklahoma, presents a study of the relative abilities of the white and colored children in the schools of that city.[4] Using the Stanford-Binet tests, eight hundred seventeen colored children and five thousand one hundred fifty-nine white children were examined by teachers, each of the latter having taken a college course in Binet testing. Four hundred eleven colored children in Grades V to XII inclusive were given the Otis Self-administering Tests. The IQ's of the groups were found to be as follows:

---

[1] *Ibid.*, p. 339.
[2] *Ibid.*, p. 399.
[3] University of Georgia *Bulletin* No. 27, December, 1926.
[4] *Elementary School Journal*, Vol. XXVI, March, 1926, pp. 542–546.

	Colored	White
Kindergarten	99	103
Grade I	93	104
Grade II	88	104
Grade III	87	102
Grade V	75	
Grade VI	74	
Grade VII	85	
Grade VIII	84	
Grade IX	81	
Grade X	84	
Grade XI	85	
Grade XII	85	
Average	82	103
Total	1228	5159

He concludes that the colored children are deficient in the abilities measured by the tests and that the relation of achievement in the two groups is about the same as the relation of their IQ's. He further finds that the idea that the colored children show slower mental growth than the whites is substantiated by the Binet test but is not in accord with the findings of the Otis test.

W. P. Davidson of Southwestern University presents arguments on both sides of the question, appears to understand the weakness of intelligence tests, but accepts Brigham's position.[1]

T. R. Garth and C. A. Whatley made a study of one thousand two hundred seventy-two children using the National Intelligence Test, Scale A, Form 1, during the fall and winter of 1924.[2] They found that the Negroes were retarded eighty-seven per cent on the average, that their mental age was below the white norms for the same grades, that the average disparity between mental age and physical age is two and one-half years, and that the IQ of southern Negro children is seventy-five. Professor Garth, in another presentation,[3] discloses the racial or national IQ's to be: White, one hundred; Chinese, ninety-eight to ninety-nine; Japanese, ninety-eight; American Negro, seventy-five; Mexican, eighty-nine. The northern Negro is much more intelligent than the southern. "The Chinaman in a

[1] *Proceedings* of the Southwestern Political and Social Science Association, 1925, pp. 7–48.

[2] *School and Society*, Vol. XXII, October 17, 1925, pp. 501–504.

[3] *Scientific Monthly*, Vol. II, September, 1926, pp. 240–245.

moralit. test was found to be more honest than Anglo-Saxons or Japanes taking the same test." He thinks that the tests used today practically eliminate the language difficulty.

D. W. Seago and T. S. Koldin present a study of Jewish and Italian children,[1] from which they conclude that the Jews are superior to the Italians and that "the difference is not essentially one of language handicap."

Dr. Bertha M. Luckey of the Board of Education of Cleveland, Ohio, summarizes the results of fourteen thousand examinations made in the last two years.[2] She found the participants to rank as follows in intelligence: Americans, Jews, Hungarians, Bohemians, Negro, Polish, Italian. She is entirely uncritical of the results.

H. L. Koch and R. Simmons, in an exhaustive study of the Negroes and Mexicans of Texas, found that in general the Negroes were inferior to the white population, both urban and rural, and the Mexicans were inferior to the Negroes in the younger age groups and superior to them in the older groups.[3] The data were collected in conjunction with the Texas Educational Survey, and it should be said that the authors are somewhat skeptical as to how large a portion of the results is due to factors other than innate intelligence.

L. Strachan, Public Schools Psychologist for Kansas City, Missouri, administered the Stanford Binet tests to Negro children in the city schools. The tests were given by trained examiners who "are kindergarten teachers engaged in teaching half of the day and in giving examinations the other half of the day." He found the Negro children inferior to the white children of the city.[4]

Miss Mercer discovered that there were differences in the color preferences of Negroes and whites but that education operates to change these and that Negroes do not have greater ability to discriminate between colors.[5]

Peterson, with the assistance of several others, presented the Rational Learning test, a mental maze test, and a disc transfer test to one hundred twenty-two twelve-year-old children, of whom seventy-four were whites and forty-eight Negroes. Five of the former and two

---

[1] *School and Society*, Vol. XXII, pp. 546–548.

[2] *Scientific Monthly*, Vol. XX, March, 1925, pp. 245–248.

[3] *Psychological Monographs*, Vol. XXXV, No. 166, 1926, pp. 1–116.

[4] *Journal of Educational Research*, Vol. XIV, October, 1926, pp. 169–177.

[5] Mercer, Florence M.: *Journal of Comparative Psychology*, Vol. V, 1925, pp. 109–120.

of the latter failed to pass the mental maze tests and were thus disqualified from further participation. In the rational learning test the whites excelled. In the disc transfers "there is a clear and reliable superiority of the whites on the time criterion." In color naming and form naming tests "the differences of speed are all in the same direction, favoring the white children."[1]

In general, then, the advocates of racial inferiority and superiority find that the races differ in intelligence, and many of them rank the races or nationalities, usually not distinguishing the two categories, according to their abilities.

The position of the middle group, those who see no evidence to warrant a conclusion one way or another, is well summarized in the statement of Professor H. C. Link. He says, "Nothing in the technique of intelligence tests, as applied so far, warrants any comparison whatsoever between the intelligence of various groups and races."[2] Burkhard,[3] for instance, quotes A. H. Arlitt to the effect that "race norms which do not take the social status factor into account are apt to be to that extent invalid," and summarizes his position with the statement that "There seems to be no definite and conclusive proof that one race surpasses another in mental capacity or native ability."

S. L. Wang[4] concludes from a study made at Ohio State University involving thirty-four Chinese, eighty-three colored men, seventy-four colored women, and forty-five Russians that the language factor cannot be ignored in attempting racial comparisons, so that failure to take it into proper consideration negates the results. C. S. Slocombe, working in the psychological laboratory of University College in London found the influence of practice to be of fundamental significance in evaluating the results of intelligence tests.[5]

Sunne expresses the conviction that it is "difficult to detect how much racial difference and how much difference in school training and social conditions contribute to the divergence in test results."[6] In another article[7] describing the results of a rather elaborate study Sunne concludes that "comparative studies show forth the complexity

[1] Peterson, J. and Others: *Journal of Comparative Psychology*, Vol. V, 1925, pp. 271–284.

[2] Board of Education Pamphlets, No. 44, p. 87.

[3] *Education*, Vol. XLVI, April, 1926, pp. 494–501.

[4] *Journal of Applied Psychology*, Vol. X, 1926, pp. 102–106.

[5] *Formal Education*, Vol. IV, 1926, pp. 173–179.

[6] *School and Society*, Vol. XIX, April, 1924, pp. 469–472.

[7] *Journal of Comparative Psychology*, Vol. V, 1925, pp. 209–220.

of mental measurements so clearly that generalizations based solely on intelligence rating must be avoided."

H. N. Bond has made a study of what he calls "non-intellectual traits" of Negro adults.[1] He finds that "if the Pressey test be a genuine measure of emotional susceptibility, the results here indicated constitute a denial of popular opinion covering this moot point." He finds little correlation between tests of intelligence and emotionality, is able to detect no distinct Negro group, and concludes that language factors and cultural isolation may easily be held to explain what differences were detected.

Professor Nutting adds the weight of an extensive experience with racial types in all parts of the world to the argument that no conclusions can yet be drawn with regard to inferiority.[2]

Dr. J. H. Oldham represents the position of the critics of the race hypothesis. In his recent book[3] he declares that it is constantly necessary "to remind ourselves that it is impossible to isolate native ability" (p. 69). He insists that "in the present state of our knowledge we have no means of determining how far observable differences are innate or how far they are due to native capacity being stimulated or hindered by circumstances" (p. 75).

Mead stresses the same point.[4] As he puts it, "A methodology adequate to deal with racial and nationality testing has not even been born" (p. 658). He classifies Ferguson's findings as a piece of circular reasoning, and insists that the three factors of language, social status and education creep into all tests, in spite of precautions.

Neifeld's position is very similar.[5] He is particularly concerned with the examinations of foreign born Americans, and his conclusion declares that the differences in scores on intelligence tests given such immigrants must be attributed to "factors related to the passage of time and unrelated to race."

Reverend James E. Gregg brings to the discussion six years experience at Hampton Institute.[6] Insisting that the army tests cannot be held to have contributed anything valid with reference to the question

---

[1] *Journal of Abnormal and Social Psychology*, Vol. XXI, 1926, pp. 267–276.

[2] *School and Society*, Vol. XXIV, July 24, pp. 89–96.

[3] "Christianity and the Race Problem." Doran, New York, 1925.

[4] *American Journal of Sociology*, Vol. XXXI, March, 1926, pp. 657–667.

[5] *American Journal of Sociology*, Vol. XXXII, November, 1926, pp. 423–432.

[6] *Scientific Monthly*, Vol. XX, March, 1925, pp. 248–254. *Southern Workman*, Vol. LIV, February, 1925, pp. 70–75.

of race differences, he presents some data gathered at Hampton in which the parties were classified according to the amount of Negro blood apparent in their color. Studying the scholastic records of the various groups, Reverend Gregg concludes that there is absolutely nothing to indicate that the presence of more or less white blood influences the achievement of the student.

Mr. Herskovits presents a rather conclusive refutation of Professor Ferguson's thesis by examining his methods of classification of Negroes and comparing the pure blood with the mixed blood groups.[1]

It may be correctly concluded that the concensus of competent scientific thought, contemplating the inability of mental testers to define intelligence, the inadequacy of all attempts to take such factors as education, social status, and language into proper consideration and the deficiencies of testing conditions, finds no proof of racial inferiority or superiority and eliminates the usual methods of determining such standing from the field of scientific usefulness.

---

[1] Herskovits, M. J.: *Pedagogical Seminary*, Vol. XXXIII, March, 1926, pp. 32–42.

THE CURVE OF RETENTION FOR POETRY[1]

PAUL L. WHITELY

Colgate University

AND

JOHN A. McGEOCH

University of Arkansas

Since the classical studies of memory made by Ebbinghaus in the early eighties, a problem of perennial interest to the student of memory is that of the shape of the curve of retention. Recent studies have modified Ebbinghaus' conclusions somewhat, and have pointed out that the shape of the curve of retention may be a function of a variety of factors. Luh,[2] for example, has demonstrated that the curve of retention varies with the method of measurement, with the degree of learning, and possibly with the age of the subjects and with the type of material. Ballard,[3] as well as Williams,[4] has shown that its shape is dependent upon the age of the subjects, and finds the presence of reminiscence more common in children. The present writers have shown[5] that the curve of retention measured by the method of narrative and interrogatory reproduction of observed material does not resemble closely the Ebbinghaus curve of forgetting, based upon the recall of nonsense syllables. They have also shown that the curve of retention varies with the method of measurement, and that forgetting is much less when measured in terms of answers to an interrogatory than when measured by narrative reproduction.[6]

Few studies have been occupied with recall over long intervals of time. That was the purpose of this study, to determine the shape of the curve of retention for poetry over long intervals of time, the

[1] The writers are indebted to Mr. Marion E. Bunch of Washington University who assisted in collecting a part of the data.

[2] The Conditions of Retention. *Psychological Monographs*, 1922, Vol. XXI.

[3] Obliviscence and Reminiscence. *British Journal Psychological Monographs Supplement*, Vol. I, 1913.

[4] A Study of the Phenomenon of Reminiscence. *Journal of Experimental Psychology*, Vol. IX, 1925, pp. 368–387.

[5] The Recall of Observed Material. *Journal of Educational Psychology*, Vol. XVII, 1926, pp. 419–425.

[6] *Cf.* Whitely and McGeoch: The Effect of One Form of Report upon Another. *American Journal of Psychology*, 1927, Vol. XXXVIII, pp. 280–284.

intervals being fifteen, thirty, sixty, ninety and one hundred twenty days.

The material used for learning consisted of six stanzas of a poem to be found on pages 28 and 29 of Foster's "Notes for Instructors" to accompany his "Experiments in Psychology." The type of poem may be discerned by quoting the first stanza. There were five other stanzas similar to this one.

> Merry it is in the good greenwood,
> When the mavis and merle are singing,
> When the deer sweeps by, and the hounds are in cry,
> And the hunter's horn is ringing.

In the college year 1924–1925, five groups of subjects, one group for each time interval, were given a fifteen, thirty, sixty, ninety and one hundred twenty day period of delay between learning and recall. In the following year the experiment was repeated, using all but the fifteen day interval. In this paper the results from the two experiments are combined. It will be noted that the conditions are so arranged that the practice effect of frequent subsequent recalls is eliminated. Thus each group recalls but twice—immediately after memorizing and again after either fifteen, thirty, sixty, ninety or one hundred twenty days, depending upon the group.

TABLE I.—MEANS OF IMMEDIATE AND DELAYED RECALLS

Group	Average number of lines	Sigma distribution	Sigma mean	Per cent delayed recall is of immediate recall
	Immediate recall			
15 day (28)	21.20	3.66	.69	
30 day (75)	22.86	2.32	.27	
60 day (65)	22.33	3.06	.38	
90 day (62)	22.46	2.87	.36	
120 day (69)	20.97	4.20	.50	
	Delayed recall			
15 day	13.56	8.51	1.61	64
30 day	10.13	7.97	.92	44
60 day	10.58	7.55	.94	47
90 day	9.31	7.51	.95	41
120 day	6.00	6.62	.79	28

TABLE II.—ERRORS

Group	Misplacements	Omissions	Additions	Substitutions	Total
Immediate recall					
15 day average	1.03	21.9	1.46	5.92	30.40
30 day average	.98	10.5	1.21	5.72	18.56
60 day average	.81	14.2	1.51	5.75	22.30
90 day average	1.12	12.1	1.56	5.67	20.40
120 day average	.94	24.3	1.53	5.72	35.40
Delayed recall					
15 day					
Average	.53	75.8	.82	5.57	82.70
Per cent delayed recall of immediate recall	51	345	56	93	272
30 day					
Average	.93	100.1	.96	5.24	107.2
Per cent delayed recall of immediate recall	93	949	79	92	580
60 day					
Average	.43	96.3	.86	4.66	102.3
Per cent delayed recall of immediate recall	53	679	57	81	459
90 day					
Average	.59	105.6	.64	5.24	112.1
Per cent delayed recall of immediate recall	53	877	41	92	549
120 day					
Average	.46	127.7	.98	3.10	132.7
Per cent delayed recall of immediate recall	49	525	64	54	373.

The experiments were conducted as group tests. The memory material, mimeographed upon one letter-size sheet of paper was placed, face-down, before each subject. When a signal was given the subject turned the page over, and memorized as much of the material as possible in the time allotted. At the expiration of fifteen minutes, time was called and the subject wrote on the back of the sheet as much of the material as could be recalled. Ample time was given for writing down all six stanzas. Then at the expiration of either fifteen, thirty, sixty, ninety or one hundred twenty days, instructions were given for the writing of the recall. After the immediate recall, the subjects in

each group were requested not to discuss the procedure of the experiment in any way in order to obviate the possibility of intercommunication among the different groups.

Both immediate and delayed recalls were scored in terms of the number of lines recalled. Note was also taken of the following errors: Misplacements, omissions, additions, and substitutions. The character of these errors may be ascertained from the connotation of the words themselves.

Since the data obtained when the experiment was repeated in the following year, were similar to those of the first experiment, the two sets of data are combined in Table I, and the curves are drawn from these combined data. The table shows the average number of lines reproduced in both immediate and delayed recall, with the sigmas of distribution and the sigmas of the means, and the ratio between the delayed recall and the immediate. The number of subjects in the combined groups is indicated in parentheses. Since five different groups are involved, and since each group gives a somewhat different mean in immediate recall, the curve of retention cannot be plotted from the differences between the means of the raw scores. Thus it will be fitting to express retention in terms of the per cent that the delayed recall is of the immediate, letting the immediate recall in each case represent one hundred per cent. A presentation of errors will be found in Table II.

Tables III and IV show the reliability of the measures. Table III presents the differences of the means for the successive intervals and their sigmas, for immediate and delayed recall. Table IV shows the cumulative differences and their sigmas. It will be observed that, according to statistical criteria, no one of the scores is significantly different from the score at the next interval. That the differences in the scores for immediate recall are not reliable is significant from our standpoint, since the differences in subsequent recall would not be explained in terms of differences in learning ability, but rather in terms of the time factor. The differences in the case of delayed recall scores, however, are not statistically significant and the differences between the thirty day and the sixty day groups and the sixty day and the ninety day groups are in no sense reliable, when viewed statistically. But with the number of subjects in each group, it would seem that it is safe to assume that the measures indicate the general tendency of the curve, at least. The consistency of this direction is certainly significant.

TABLE III.—DIFFERENCES AND THEIR SIGMAS—IMMEDIATE AND DELAYED

Interval	Mean	Difference	Sigma of difference
		Immediate recall	
15	21.20	1.66	.74
30	22.86		
30	22.86	.53	.34
60	22.33		
60	22.33	.13	.52
90	22.46		
90	22.46	1.49	.61
120	20.97		
		Delayed recall	
15	13.56	3.43	1.85
30	10.13		
30	10.13	.45	1.31
60	10.58		
60	10.58	1.29	1.31
90	9.31		
90	9.31	3.31	1.20
120	6.00		

TABLE IV.—CUMULATIVE DIFFERENCES AND THEIR SIGMAS

Interval	Mean	Difference	Sigma of difference	Interval	Mean	Difference	Sigma of difference
15	13.56	3.43	1.85	30	10.13	.82	1.34
30	10.13	....	....	90	9.31		
15	13.56	2.98	1.86	30	10.13	3.51	1.21
60	10.58	....	....	120	6.62		
15	13.56	4.25	1.87	60	10.58	1.27	1.33
90	9.31	....	....	90	9.31		
15	13.56	6.94	1.79	60	10.58	3.96	1.22
120	6.62	....	....	120	6.62		
30	10.13	.45	1.31	90	9.31	2.69	1.23
60	10.58	....	....	120	6.62		

Moreover, if points remote from each other on the curve are validly different, the general shape of the curve as here determined is essentially reliable. In order to discover this, the cumulative differences were computed between each interval and all following points, as shown in Table IV.

We shall deal first with the influence of the time intervals upon correct recall. The curve of retention, as shown in Fig. 1, falls rather abruptly from immediate recall to recall after thirty days. There is a slight negative acceleration at sixty days, dropping slightly at ninety days and slightly more abruptly at one hundred twenty days. In other words, after fifteen days, thirty-six per cent of the material has been forgotten. And after thirty days, fifty-six per cent of the material was lost, or a little over half. Using nonsense syllables, Ebbinghaus[1] found that practically four-fifths of the material had been forgotten after the same interval. If a comparison is permitted, we may conclude that the retention of sense material, which in this case has the advantage of rhythm and rhyme, is considerably greater after a period of thirty days than the retention in the case of nonsense material. Our curve of retention indicates that the amount of forgetting of poetry from thirty days to ninety days is negligible. Radossawljevitch[2] found that the forgetting of nonsense syllables was practically complete after one hundred twenty days. For poetry we find a retention of twenty-eight per cent after one hundred twenty days. The slight reminiscence at sixty days, as compared with the thirty day interval, is probably due to some chance factor, and therefore is not significant. The only safe conclusion, it would seem, is that forgetting of poetry after thirty days increases very gradually.

The shape of this curve differs from the type of curve obtained by the writers in a previous experiment.[3] Measured by narrative reproduction, we found that the curve of retention for observed material falls very nearly in a straight line from immediate recall to recall after ninety days, showing slight negative acceleration at one hundred twenty days. Our curve in this study more closely approximates in form the classical Ebbinghaus curve, but the time intervals are, of course, longer.

---

[1] "Uber das Gedachtniss; Untersuchungen zur experimentellen psychologie." Leipsig, 1885.

[2] "Das Behalten und Vergessen bei Kindern und Erwachsen." Leipsig, 1907.

[3] The Recall of Observed Material. *Journal of Educational Psychology*, Vol. XVII, 1926, pp. 419–425.

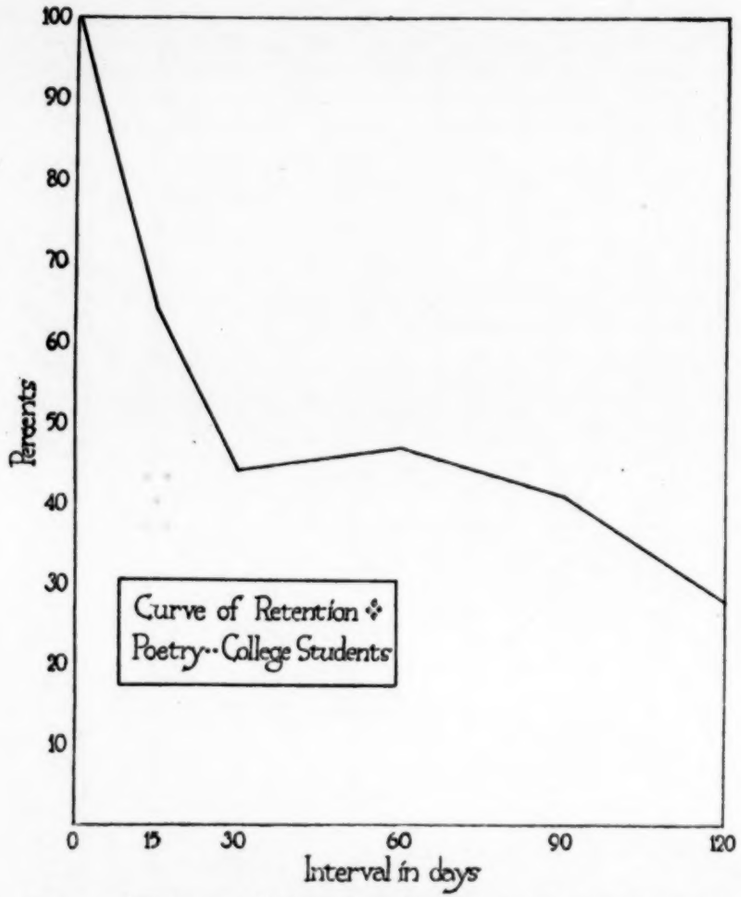

FIG. 1.—The curve falls rather abruptly from immediate recall to recall after thirty days. There is a slight negative acceleration at sixty days, dropping slightly at ninety days and slightly more abruptly at one hundred twenty days. It closely approximates in form the classical Ebbinghaus curve.

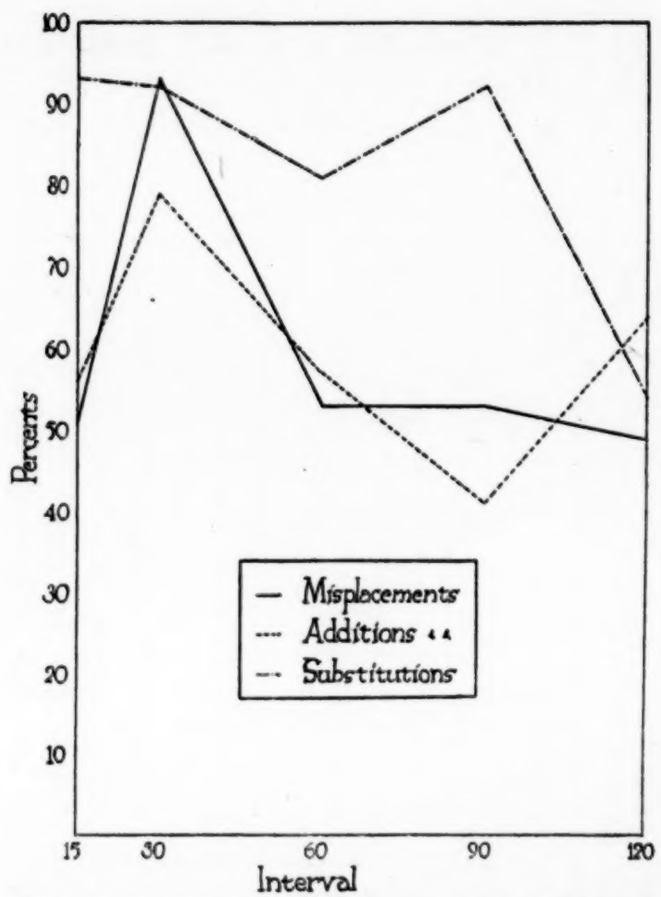

FIG. 2.—The error curves show a fairly uniform increase in errors at thirty days, and a decrease after thirty days. They, are, however, somewhat irregular.

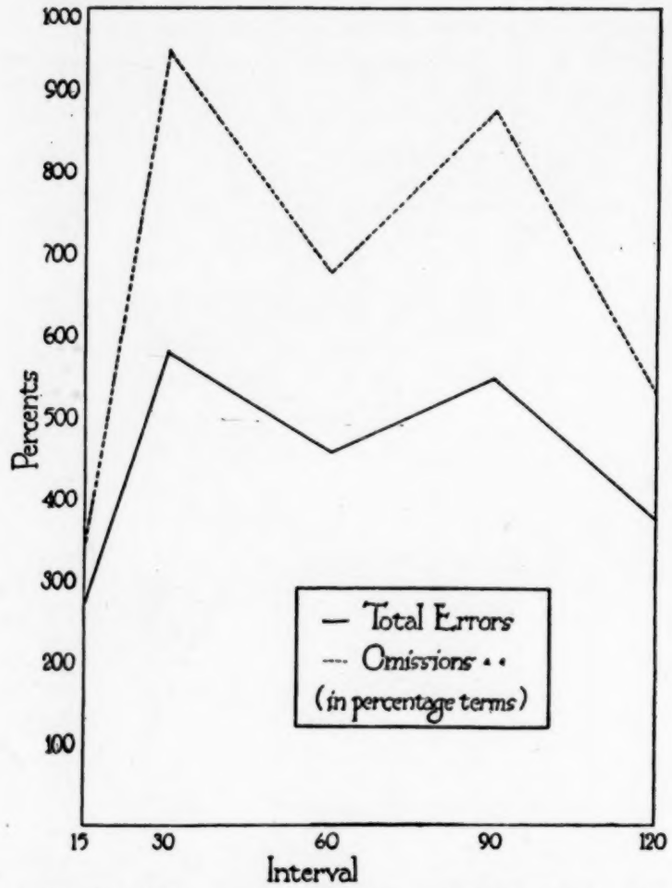

FIG. 3.—The curve showing omissions rises quite abruptly from fifteen to thirty days, and decreases more gradually after thirty days. The same general tendency is indicated for total errors. Again they are characterized by irregularity, however.

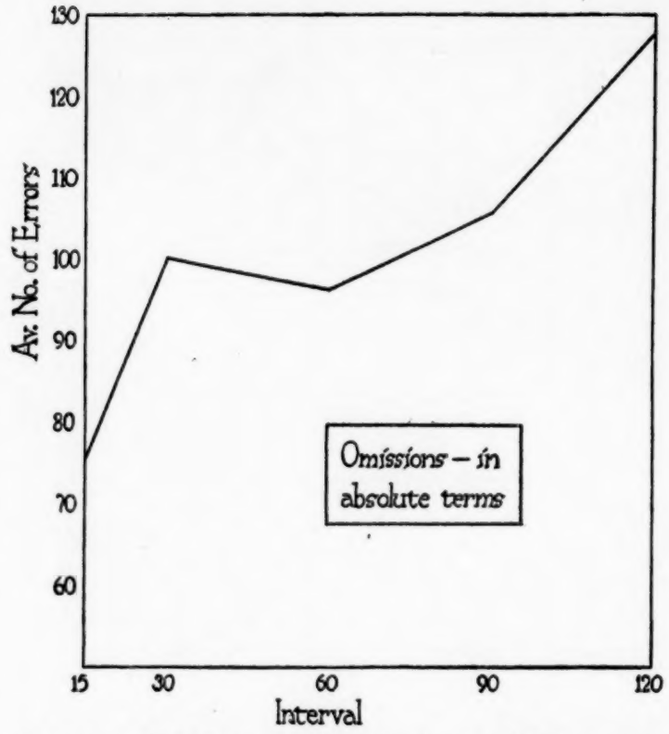

FIG. 4.—Omissions in absolute terms exhibits a curve that rises quite abruptly from fifteen days to thirty days, more gradually from thirty days to ninety days, and fairly abruptly from ninety days to one hundred twenty days.

The error curves, found in Figs. 2, 3, and 4, are irregular and more difficult to interpret. Perhaps the most striking feature about these curves is the fairly uniform increase in errors at thirty days. It is conceivable that after fifteen days, recall was sufficiently good so that the subject did not attempt to write down items which he did not remember fairly well. That is, his recognition memory was sufficiently good so that he knew whether the item was right or wrong, and if wrong did not write it down. After thirty days, however, the memory "traces" had become sufficiently weakened so that recognition memory was wavering and uncertain. Consequently, the subject was more willing to take the risk of writing items of which he was less certain than when recognition was better. And after sixty days the subject has probably forgotten all that had not been well learned, or probably over-learned; hence there was not the provocation to write down items of which he was not fairly certain. From an examination of the curves, this explanation would seem to apply more particularly to misplacements and additions. Omissions, as represented in terms of absolute number of omissions, exhibits a curve that rises quite abruptly from fifteen days to thirty days, remains roughly constant from thirty to sixty days, and rises quite rapidly from sixty days to one hundred twenty days. It is more scientifically correct to represent omissions in percentage terms, however, as in Fig. 3, which is extremely irregular and difficult to analyze.

Conclusions

1. The curve of retention for poetry over long intervals of time, falls quite abruptly from immediate recall to recall after thirty days, and very gradually from thirty days to one hundred twenty days. The drop from thirty days to ninety days is relatively insignificant.

2. The form of the curve of retention for poetry approximates more closely the Ebbinghause curve of retention than the curve obtained by the authors for the retention of observed material in a previous experiment.

3. In general, errors increase from fifteen days, to thirty days, and decrease after thirty days. A possible explanation is suggested.

4. Omissions in terms of the average number of errors gives a curve that rises rather abruptly from fifteen days to thirty days, and gradually from thirty days to one hundred twenty days. In terms of the ratio of delayed recall to immediate, the curve representing omissions is quite irregular.

THE RELATIVE INFLUENCE OF TWO TYPES OF MOTIVATION ON IMPROVEMENT

VERNER MARTIN SIMS

Louisiana Polytechnic Institute, Ruston, La.

The experiments here reported are attempts to evaluate the influence upon improvement of two different types of motivation: in the first type, which we have called individual-motivation, the individual competes against his own record and against that of other individuals of like ability; in the second type, which we have called group-motivation, the individual as a member of a group, competes against another group. Two experiments are reported, the first using substitution as the function to be improved, the second using rate of reading.

The method of conducting the experiment in substitution and the results of the experiment are as follows:

An initial practice period of three minutes at substituting digits for letters was given to one hundred twenty-six college sophomores and juniors, and on the basis of the number of substitutions made per minute, three sections of twelve each were equated by selecting trios the members of which made the same initial score, one member of the trio going into each of the three sections.[1] This manner of selection made the sections equal in range as well as in central tendency.

These three sections practiced at substitution three times a week for a total of twelve practice periods, using the same as that used in the initial period, but with varying forms. The practice periods after the initial one were two minutes in length.

The three sections were motivated as follows:

Section I. The Control Section.—No motivation other than that which came incidentally from seeing their own progress and that of their neighbors was used. The blanks were collected immediately after the practice, no attempt was made to prevent them from watching their progress or to encourage them to improve. They perhaps knew approximately what their own score was; and they left the room in a body immediately after the practice, perhaps discussing it to a slight extent.

Section II. The Group-motivated Section.—The section was divided into two groups, whose total initial score was approximately equal, and they were competed against each other. Before beginning practice the average score of each group for the preceding practice period was read to the section, the amount of

[1] Sixteen such trios were originally formed, but only twelve completed the series of practices.

improvement for each group was given, and a graph showing the total progress of each group was presented. They were then encouraged to put forth their best efforts to make their respective groups win. After the practice the blanks were collected and the groups dismissed.

Section III. The Individually-motivated Section.—The section was divided into pairs, the two members of a pair having approximately equal initial ability. Before beginning practice each individual score was called out, and as they were called each member entered on a graph his own score and that of his particular competitor. In addition, the names of the three persons who showed the greatest improvement with the three that made the least improvement and the names of the three persons who made the highest scores with those three that made the lowest scores on the previous day's practice were reported to the class with the praise or blame which the record deserved. All were encouraged to surpass their previous records and if possible to defeat their respective competitors. Immediately after the practice the blanks were collected and the groups dismissed.

In all three sections the attempt was made to keep the personal enthusiasm of the experimentor out of the procedure. The results were given in a more or less stereotyped manner, and the conversation was limited to the giving of the information mentioned above. None of the members knew exactly the nature of the experiment, although they were perhaps conscious of the fact that there were different groups practicing at substitution under different conditions.

The average results for each section, expressed in terms of the number of substitutions made per minute, for the twelve practice periods are shown in Table I.

TABLE I.—THE AVERAGE NUMBER OF SUBSTITUTIONS PER MINUTE FOR THE THREE GROUPS

	First	Second	Third	Fourth	Fifth	Sixth	Seventh	Eighth	Ninth	Tenth	Eleventh	Twelfth
Section I.....	36.0	42.8	51.4	55.3	60.0	64.8	69.8	70.3	68.4	72.9	74.8	70.8
Section II.....	36.1	39.4	46.2	53.1	55.8	58.1	62.3	66.4	70.4	72.3	76.9	74.7
Section III.....	36.2	41.7	51.8	60.4	66.0	70.3	75.8	79.4	83.6	89.5	91.3	95.4

Using the average of the last two practice periods as the final ability we find that:

Section I improved from 36 substitutions per minute, to 72.8 substitutions per minute, or 102.2 per cent.

Section II improved from 36.1 substitutions per minute, to 75.8 substitutions per minute, or 109.9 per cent.

Section III improved from 36.2 substitutions per minute, to 93.3 substitutions per minute, or 157.7 per cent.

The method of conducting, the experiment in the rate of reading and the results are as follows:

Seventy-six college sophomores and juniors were allowed to work for two minutes and forty seconds on the Monroe Standardized Silent Reading Test, Form 2, and on the basis of the words read per minute three sections of fifteen each were equated by selecting trios with approximately equal initial scores, one member of each trio going into each of the three sections. Thus, as in the first experiment, the groups were of the same range as well as central tendency.

These three sections practiced at reading for three minutes three times a week for a total of twelve practice periods. The reading practice consisted of (1) reading from Gates' Psychology for Students of Education" and at the same time answering simple Yes-No questions concerning the material read;[1] (2) reading material from the same book at the fastest rate possible yet with understanding (they being sole judge of this understanding); (3) working for three minutes on some reading test other than Monroe's. These different forms of practice were presented on alternate days in a cycle. That is, each section had found practice periods on each of the above types of material: the first, fourth, seventh and tenth practices being devoted to the first type, the second, fifth, eighth and eleventh to the second type and the third, sixth, ninth and twelfth to the third type of material. At the conclusion of this series of practices they were again allowed to work for two minutes and forty seconds on the Monroe Reading Test, this time using Form 3.

The three sections were motivated as follows:

Section I. Control Section.—At the beginning of the first practice they were urged to try to improve their rate of reading, but they knew nothing of their scores, neither on the initial test nor on the practice material.

Section II. The Group-motivated Section.—The total section was divided into two groups of approximately equal ability and they were competed against one another. Before beginning practice they were given the average score of each group for the previous practice period, and the last average score made by each group on the particular type of material on which they were about to practice. In addition, they were presented with graphs showing the progress of each group in the type of material with which they were that day to work. They were urged each day to do everything possible to make their own group win.

Section III. Individually-motivated Section.—As before, the members were paired; each individual keeping his own record, and that of another member whose

[1] These questions were presented on mimeographed blanks, and included more material than the fastest reader could cover in the alloted time.

initial score was approximately equal to his own for each of the three types of material used in the practices. In addition to reading the previous day's scores, the names of the three high ranking persons and the three low ranking persons for the previous practice with similar material were given. Each member of the section was urged to make an effort to surpass his own record and that of his particular competitor.

The average initial and final scores on the Monroe Tests expressed in terms of the number of words read per minute, are shown in Table II.

TABLE II.—INITIAL AND FINAL RATE OF READING FOR THREE SECTIONS

	Initial rate (words per minute)	Final rate (words per minute)	Per cent improvement
Section I	167.3	181.9	8.7
Section II	167.5	191.9	14.5
Section III	167.7	226.0	34.7

The norms for Form 3 of the Monroe Test, as given by the author, for the various grades of the elementary school are slightly higher than those for Form 2, consequently we may conclude that it is the easier of the two forms and the increased rate of Section I represents very slight if any improvement. The results are to be considered as relative measures of improvement rather than indications of absolute improvement in reading rate.

Those familiar with the Monroe Test know that it consists of a series of paragraphs each followed by a multiple-choice question concerning its contents. Since the test forms are intended for use with grade children, none of the subjects participating in this experiment should have difficulty in comprehending the meaning of any of the paragraphs; consequently, the percentage of the questions attempted that are correct would be a rough measure of the efficiency of reading. We can thus check the effect of the increased speed upon efficiency. Table III presents the initial and final efficiency of each section, this efficiency being expressed in terms of the percentage of those questions attempted that were answered correctly.

The increased rate of reading, then, has tended to improve the efficiency of Sections I and III, while Sections II has a lower average efficiency on the final test. These differences, however, are so slight that they cannot be interpreted as meaning anything other than the fact that the efficiency has not been materially affected by the practice.

TABLE III.—THE AVERAGE EFFICIENCY OF THE THREE SECTIONS ON THE INITIAL AND FINAL READING TESTS

	Initial test			Final test			
	Number tried	Number right	Per cent of efficiency	Number tried	Number right	Per cent of efficiency	Per cent change
Section I.....	8.3	7.4	89.2	10.3	9.6	93.2	4.0
Section II.....	8.1	7.5	92.5	10.5	9.5	90.5	−2.0
Section III.....	8.2	7.5	91.4	12.4	11.9	95.9	4.5

To summarize, with these two types of material, rate of reading and rate of substituting, and with the groups here used, individual-motivation is vastly superior to group-motivation and group-motivation is but slightly superior to no motivation other than that which comes incidentally in learning. To the extent that reading and substitution are typical of learning in general one may say that for the groups here concerned individual-motivation is the superior form of motivation. To the extent that these students are typical of students in general, one may say that for improving the rate of substituting and the rate of reading individual-motivation is superior to group-motivation.

From the standpoint of efficiency in learning, there is urgent need for a repetition of these experiments using different functions for improvement and different age groups as subjects. More than this, it is essential that the groups be tested after periods of no practice in order that the permanency of the improvement may be ascertained. The superiority of the individually-motivated group may represent a "forced growth" which in time will disappear. And finally, before any practical educational value can be derived from such information, we must know that a function thus improved will show general improvement in the various situations where it may be used, *i.e.*, that an increased rate of reading under drill conditions such as those here outlined will mean an increased rate in other situations involving reading.

LEARNING AND RETENTION OF LATIN WORDS AND PHRASES

J. P. ANDERSON

AND

A. M. JORDAN

The University of North Carolina, Chapel Hill, N. C.

The present investigation attempts to study certain phenomena of the memory and retention of English equivalents of Latin words. We can probably best state our aim in the form of certain questions which seemed to us to be proper for study and upon which the present investigation casts some light. These are:

1. What is the rate of forgetting of the English equivalents of Latin words?
2. How does the rate of memorizing correlate with the rate of forgetting among pupils?
3. How does the ability at memorizing and retaining English equivalents of Latin words correlate with (*a*) intelligence as measured by standard intelligence tests, (*b*) with word knowledge, and (*c*) with ability in silent reading?
4. What kinds of Latin-English word pairs are most readily learned and longer retained?
5. How effective is the teaching of English-Latin derivation upon the retention of English equivalents of Latin words?
6. Are words which are learned in groups (idioms and phrases) more easily learned and readily retained than words learned singly?

The subjects were the pupils of Grade VII of the Chapel Hill (N. C.) Public School. The average age of the group was twelve years and eleven months. The pupils were thirty-one in number and full records were obtained for thirty. None of the subjects had, prior to the investigation, any acquaintance with the Latin language. Intelligence quotients were obtained for twenty-three of the group from scores on the National Intelligence Tests A-1 and B-1, given the previous year. Scores were secured on the Thorndike Test for Word Knowledge and on the Thorndike-McCall Silent Reading Ability Test for thirty of the subjects.

The material used for study consists of approximately two hundred fifty words taken from the list of the five hundred most common words of Cæsar compiled by Lodge. From this list were selected the nouns, adjectives, prepositions and conjunctions. It was intended originally to study the entire five hundred words but the time required for the

recall tests was such that only half the words could be presented in the time allotted to the experiment. The words are grouped in vocabularies. Each vocabulary represents the material studied in one daily thirty minute period. The words are divided into three groups differing in their similarity of sound to their English equivalents. Thus in the first group of words, the English equivalents are almost identical in sound with the Latin words, for example, *provincia*—province. These words are called "identical words" in this paper. All the words in this group are not so identical in sound as the one cited for example, *mater*—mother is placed in this group. There are thirty-one "identical" words. The first test contained the following words:

provincia	province
causa	cause
natura	nature
memoria	memory
fortuna	fortune
victoria	victory
ancora	anchor

The next list contains words whose English and Latin sounds are dissimilar but for which there are derivative English words closely associated to the Latin word in sound. In the vocabulary lists the derivative words are placed to the right of the English equivalent for the pupil to use in studying if he so desires. The use of the derivative word in study was left to the choice of the pupil. We will call this group of Latin words "association words." There are seventy-five of this type given in the study. The first test contained the following words:

fuga	flight	fugitive—one who flees
lingua	language	lingo—slang for language
copia	abundance	copious—large
amicitia	friendship	amicable—friendly
filia	daughter	filial—filial devotion
iniuria	injustice	injury—i for j
via	street, way	via—by way of
gratia	favor	"grat" slang for favor
vigilia	a night watch	vigil—while angels their vigils keep
iustitia	justice	justice—i for j

The method of dealing with these words will be explained more fully later.

The third group of words are those between whose Latin and English form there is no sound similarity or derivative connection familiar to the pupil and which depend for their learning on rote memory. We have called these words "non-association words." There are ninety of these words. The first test contained the following words:

telum	javeline
fossa	ditch
angustia	narrow pass
contumelia	insult, abuse
lacrima	tear
insidia	treachery
mora	delay
cibus	food

In the fourth place, certain Latin phrases were presented with their English equivalents. Examples of these phrases are:

(1) *e pluribus unum*—from many, one
(2) *Sic semper tyrannis*—so (be it) ever with tyrants
(3) *vox populi vox Dei*—the voice (of the) people (is the) voice of God.

As we were to study the advantage that might result in learning English equivalents of Latin words when there were derivative words given it was necessary to know whether the derivative word which we had supplied was intelligible to the child. Thus in the pair *annus*—year, the derivative word being "annual" if the child does not know the word annual the derivative word is not an aid and is likely to be a hindrance. In this manner all the derivative words used were given to the subject in a test for English meaning. Of the seventy-five derivatives which we had supplied, it was found that only forty were known to a majority of the pupils. The criterion used in classing the words of this group as "association" or "non-association" word was that fifty per cent of the pupils were familiar with the derivative word given. The thirty-five words whose derivative words was not known to fifty per cent of the group were therefore classed as "non-association words."

Procedure

The pupils were tested on each school day for a period of five weeks with the exception of three days near the end of the period when school was suspended. Two more recall tests were given after eight weeks—that is, five weeks after the testing period closed. The total number

of days on which tests were given was twenty-four. The daily period was thirty minutes, coming at 12:15 to 12:45, the period preceding the noon recess. Tests for recall were given as follows: Immediately after study, on the following day, one week after study, three weeks after study, and eight weeks after study.

One vocabulary was given each day. Each pupil was furnished with a typewritten copy on which was written with their appropriate meaning seven identical words, ten associative words, and eight non-associative words. These vocabulary sheets were kept face down after being distributed until on the word of the examiner the pupils began studying. The time allotted for study amounted to fifteen seconds per word. For a vocabulary of twenty-five words (which was the length more commonly used) the study time was six and one-fourth minutes. This time limit, which was adhered to strictly throughout the experiment, was found to be sufficient for the upper fifty per cent of the pupil group. That is, it was sufficient in that a majority was generally able after studying for this length of time to reproduce immediately all the words of the tests given. The results on the tests for immediate reproduction show that the upper fifty per cent group reproduced about ninety per cent of the words studied. The possible advantage to be gained for a longer study period to the lower fifty per cent of the group was believed to be overbalanced by the consideration that those who learn the words before the time is up become restless, while the shorter time limit keeps all the group working at a high level. At the end of the study period the pupils were instructed to fold the vocabulary sheet so that the Latin words only were visible and on another sheet of paper to write the English equivalents. Five minutes were generally allotted for this operation although it was desired to give sufficient time for the slowest to write all they remembered. After writing all the words remembered the pupils were allowed to go over the list of words and note their errors and omissions but to make no addition to the first copy. They wrote the words gained after correction on a separate sheet of paper and these words are not counted in the totals given in the tables. This method of acquainting the pupils with their degree of success or error was thought necessary for eliciting interest and as such was found to be successful.

The same order of words in the list was used in the tests for recall on the following day but for the other recall tests, *i.e.*, recall after one week, three weeks, and eight weeks the Latin words were presented in a different order from that in which they were learned.

Results

In answer to the questions: "What is the rate of forgetting English equivalents of Latin words?" we present the data of Table I, and Fig. 1 which show curves of forgetting plotted from these tables.

Table I gives the average recall for the group on all the recall tests given—recall immediately after learning, recall after one week, recall after three weeks, and recall after eight weeks. The records

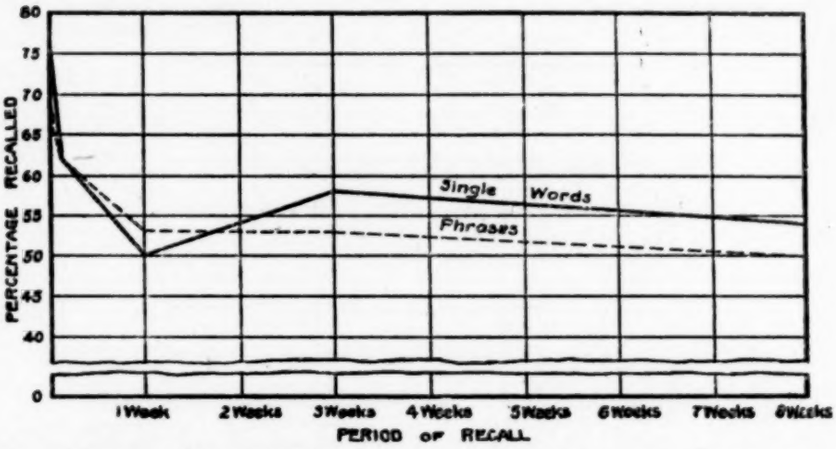

Fig. 1.—Curves of forgetting for Latin words and phrases. Tests were given on the same day of study, on the next day after study, after one week, after three weeks, and after eight weeks.

are based on one hundred sixty words which were presented in the first seven vocabularies.

Table I.—Per Cent of Retention of Latin Words and Phrases

	Words	Phrases
Immediate recall	76	69
Recall on following day	62	62
Recall after one week	50	53
Recall after three weeks	58	
Recall after eight weeks	54	50

In the curve of forgetting with nonsense syllables from data by Ebbinghaus, where the nonsense series is learned to the point of one successful reproduction, it is seen that four-fifths of the loss occurs during the first day. The curve obtained from data by Radossawljewitsch, when the nonsense series is learned to the point of two successful repetitions is more similar to those for poetry and Latin-English pairs. Here the first day's loss is about one-half that of the first three weeks', but the curve does not come to a level during the entire period of the test (thirty days), whereas our curves show no

loss from the first to the eighth week, in fact they show a slight gain. This is readily explicable along two lines: Each attempted recall at periods of one day, one week, two weeks, and three weeks would tend to strengthen the connections; and the material was so organized that a Latin word or phrase was always paired with an English equivalent and thus in many cases the Latin word would give a direct clue to the English equivalent. Possibly the more significant finding is that

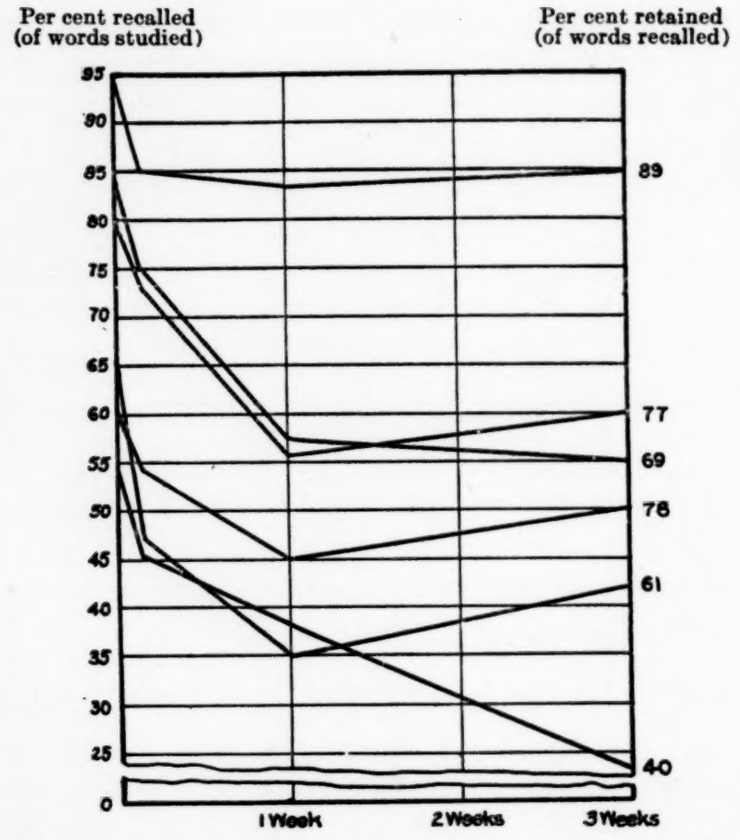

FIG. 2.—Curves of forgetting for six groups of pupils, five in each group. Note that in two cases the groups maintain their relative positions at the end; in four, they do not.

after a period of two months only one-half of the meanings of Latin words was retained.

Secondly, what is the relation between the rate of memorizing and the rate of forgetting? Do those individuals who learn more words within the six minutes as reproduced on the immediate test retain more after an interval of eight weeks? The results obtained with the thirty pupils studied here indicate that those who learn most remember most, that those who learn least remember least, etc.

Figure 2 shows that the group which learned most words retained a greater per cent than any other group; and the group which learned fewest words retained the smallest per cent of them. This is true of each of the ten individuals in the two groups as well as of the average of the groups. However, among the four middle groups there is not this exact inverse relationship between amount learned and per cent remembered. As can be seen from the per cents of recall on the right hand margin of Fig. 2, group four which is second in per cent of delayed recall is fifth in immediate memory. However the differences in per cents of recall and in memory among the four middle groups are small. Considering the six groups together we can say that good retention goes with rapid memory rather than the opposite. To express this relation in another way, the coefficient of correlation between initial retention and final retention eight weeks later in the case of one hundred sixty words is .79 $\pm$ .07. The corresponding coefficient with the second test three weeks removed was .75 $\pm$ 0.7. These results correspond with results obtained by Gates and Thorndike (see bibliography).

Thirdly, how does the ability to memorize and retain Latin-English equivalents correlate with intelligence, with silent reading, and with word knowledge? We find the following coefficients of correlation:

TABLE II

	Intelligence quotients (N. I. T.)		Silent reading		Word knowledge	
	r	PE	r	PE	r	PE
Immediate memory.....	.65	.12	.46	.14	.20	.18
Recall after eight weeks.....	.58	.14	.468	.14	.208	.18

The correlations of Latin-English word memory with intelligence are high; with silent reading ability are significant, and with word knowledge are low.

Fourthly, what kinds of Latin-English word pairs are most readily learned and retained? We have already indicated the three types of Latin-English pairs which were studied: Identical words, associative words, and non-associative words. The per cents of reproduction and delayed recall for each type of words is as follows (Table II):

TABLE III.—PERCENTAGE RETAINED OF DIFFERENT TYPES OF LATIN WORDS

	Immediate reproduction	Recall one week	Recall three weeks	Recall eight weeks
Identical words	93	80	90	90
Associative words	88	58	56	65
Non-associative words	66	48	39	37

It is clear that the identical words come first, then the associative words and finally the non-associative words. One will notice that at the end of one week the non-associative words were quite close to the associative words, the difference in percentage being ten; while at eight weeks the difference is twenty-eight.

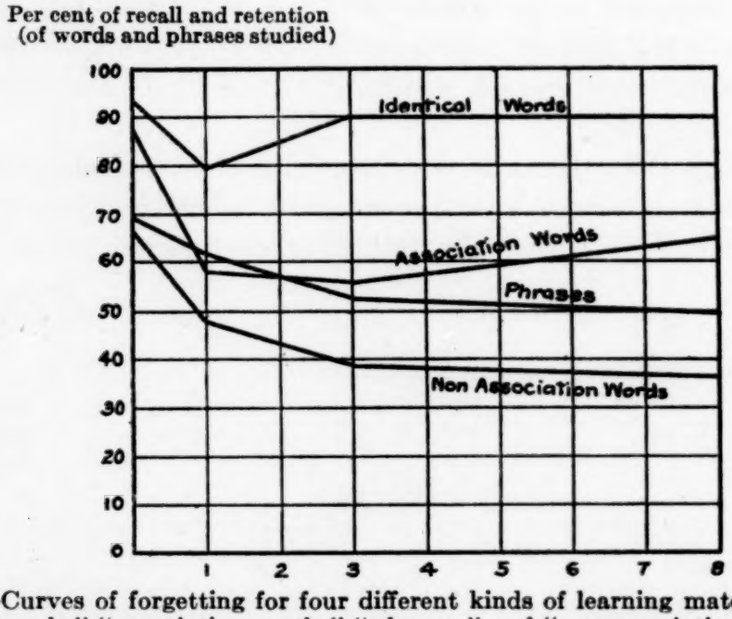

FIG. 3.—Curves of forgetting for four different kinds of learning material: "identical words," "association words," "phrases," and "non-association words."

The curves of forgetting for the three types of words studied and forty idioms are given in Fig. 3. There were thirty words of the identical type, forty associative, and ninety non-associative words. For comparative purposes instead of using the entire ninety non-associative words we used the forty most commonly recalled ones. We judge that the most evident factors influencing learning are length of word, and the meaning of the word, *i.e.*, whether it is noun, the name of familiar object of abstract quantity, adjective, or pronoun. So in choosing the forty most favorable words from a list of ninety studied

and scored we can be reasonably certain of not weighting the dice in favor of an advantage that might accrue to the words taught by derivation. Again, comparing the words of the two types with respect to the attributes noted above (length and meaning) we find them as nearly equivalent as is demanded for comparison.

The curves shown on Fig. 3 are sufficient to indicate the results. The drop in the curve of forgetting of the "identical" words is partly explainable by the confusion in the minds of the pupils produced by an inverse order of words on the test list, but we would attribute the rise in this curve to some extent (and in the associational word curve to a greater extent) to the fact that the pupils had begun to look for identical and derivative meaning in some of the words. In these latter tests especially we find many words given meanings as identical words which were not identical and some erroneous associations formed. Thus it might be pointed out that while the derivative devices were favorable to learning those words in which derivatives were legitimate, they acted as a handicap toward those of a non-derivative nature.

Additional facts bearing on the question of the value of teaching derivation as a means toward teaching Latin vocabulary (more interesting, probably, than valuable) are answers obtained from the pupils on a questionnaire. Questions asked were: "Have you used the derivative words supplied in learning the Latin words?" "Do you think the derivative words help you to memorize the Latin words?" and "Did you 'make-up' any associational words where there were none given?" All but two pupils think that the derivative words help them. Some say that "some times they help and sometimes they hurt." One or two say that they do not pay any attention to them usually. In answer to the third question about half of the class assert that they think of associational words where none are indicated. One girl said: "I just happened to think of some." Several asked toward the end of the period where certain English words came from and if we do not get such and such a word from this Latin word. This attitude of looking for derivatives is valuable in itself, we think.

Other responses made to some of the words given may be of interest. Thus we gave *voluntas*—wish: Derivative word—voluntary—"of one's own wish." The equivalent given to *voluntas* was "voluntary" as often as the correct response "wish." We gave the word *fuga*—flight: Derivative word fugitive—"one who flees." Here we explained that a fugitive was one who runs away, such as a thief trying to escape.

Answers to the word *fuga* were—"thief," "to flee" and "flight." Another word was *vigilia*—a night watch: Derivative word—vigil. The explanation given was that vigils were watches as "angels their vigils keep." Responses to *vigilia* were—"night watchman," "angel," and "virgin." It is evident to us that, for the presentation of derivatives to be of advantage to the learning of the Latin word, it is necessary that the pupil be conversant with such derivatives, otherwise the derivatives instead of facilitating the connection between the Latin and English equivalents only provide an additional connection to be formed.

The last question which we proposed to answer in this study is whether words in groups as idioms and phrases are more readily learned and longer retained than words studied singly. The average per cents of reproduction and recall for the forty idioms and phrases studied are given in Figs. 1 and 3. As the curves indicate, the rates of memory and retention for this material is less than that for association words and greater than that for non-association words. Certain conditions of memory of this material should be noted. The pupils, as mentioned in the beginning of this study, had no knowledge of grammatical construction or inflection of Latin words. This fact minimized the advantage to be gained from learning words in idioms and phrases since the pupil was apt to see the phrase as a group of isolated words than as a connected phrase. The method of scoring the phrase as right or wrong was to give no credit if any of the words were wrong. However, the pupils were instructed to put down on their papers the meaning of any single words that they knew in the phrase. Among the papers of fully a third of the pupils, there were five or more of these disconnected words indicated correctly.

Our interpretation of the data is that there is advantage to be gained from teaching Latin-English equivalents by presenting the material in this form though not so great an advantage as we believed before the experiment was conducted. The extent of this advantage is not certain from this study.

Summary

1. In this study, the curve of forgetting for English equivalents of Latin words is less abrupt than that for nonsense material and is similar to that for English poetry.

2. There is a high correlation between immediate memory and delayed recall, *i.e.*, the order of pupils on the immediate memory

tests, is generally the same as that on the delayed recall tests. To put it another way, those who learn more, retain more. With the extreme cases (the five best and five poorest pupils) those who learn most lose a smaller per cent of words learned and those who learn fewest lose a larger per cent of words learned. Among the middle twenty cases this relationship does not obtain.

3. (*a*) There is a high correlation between intelligence and immediate memory of Latin-English equivalents and between intelligence and delayed memory of Latin-English equivalents.

(*b*) There is a significant correlation between silent reading ability and immediate memory of Latin-English equivalents and between silent reading ability and delayed memory of Latin-English equivalents.

(*c*) The correlation found between word knowledge and immediate memory of Latin-English equivalents and between word knowledge and delayed recall of Latin-English equivalents is not great enough to be significant. This finding is subject to question in the light of contrary findings by other investigators.

4. The per cent of material reproduced on immediate reproduction tests and the per cent of recall on eight weeks recall tests shows that the materials studied stand in the following order with respect to the per cent of reproduction and the per cent of recall: (1) Identical words; (2) associative words; (3) idioms and phrases; (4) non-associative words.

5. This experiment and other experiments cited indicate that the presentation of Latin-English derivation is effective in the memory and retention of Latin-English equivalents.

6. The findings for immediate memory and retention of Latin idioms and phrases, though inconclusive, indicate an advantage to be gained in teaching Latin vocabulary by this means.

References

(*a*) Latin

1. Allen, W. S.: "A Study in Latin Prognosis." Columbia University, New York City, 1924.
2. Hamblen, A. A.: "An Investigation to Determine the Extent to Which the Effect of a Study of Latin upon a Knowledge of English Derivatives Can Be Increased by Conscious Adaptation of Content and Method to an Attainment of this Objective." University of Pennsylvania, Philadelphia, 1925.
3. Haskell, R. I.: "A Statistical Study of the Results of Teaching Derivation in the Ninth Grade Latin Classes and in the Ninth Grade English Classes in Four Philadelphia High Schools." Philadelphia, 1923.

4. Lodge, Gonzalez: "The Vocabulary of High School Latin." Columbia University, New York, 1922.
5. Meek, Lois H.: "A Study of Learning and Retention in Young Children." Columbia University, New York, 1925.
6. Thorndike, E. L.: "Educational Psychology." Vol. II, New York, 1925.
7. Schmidt, Austin G.: "The Effect of Objective Presentation on the Learning and Retention of Latin Vocabulary." Loyola Univ. Press, Chicago, 1923.

(*b*) Psychology of Learning

1. Book, W. F.: The Role of the Teacher in the Most Expeditious and Economic Learning. *Journal of Educational Psychology*, Vol. I, pp. 183–189.
2. Dearborn, W. F. and Brewer, J. M.: Methods and Results of a Class Experiment in Learning. *Journal of Educational Psychology*, Vol. IX, pp. 62–82.
3. Gates, Arthur I.: Correlation of Immediate and Delayed Recall. *Journal of Educational Psychology*, Vol. IX, pp. 489–496.
4. Henaemon, V. A. C.: The Relation between Learning and Retention and the Amount to Be Learned. *Journal of Educational Psychology*, Vol. II, pp. 476–484.
5. Lyon, D. O.: The Relation of Length of Material to Time Taken for Learning and Optimum Distribution of Time. *Journal of Educational Psychology*, Vol. IX, pp. 1–9; 89–91; 155–163.
6. Myers, G. C.: Recall in Relation to Retention. *Journal of Educational Psychology*, Vol. V, pp. 119–130.
7. Myers, G. C.: Some Correlations between Learning and Recall. *Journal Educational Psychology*, Vol. VII, pp. 546–547.
8. Perkins, N. L.: Value of Distributed Repetitions in Total Learning. *British Journal of Psychology*, Vol. VII, pp. 253–261.
9. Pyle, W. H.: One Function of the Teacher in Memory Work. *Journal Educational Psychology*, Vol. I, pp. 474–476.
10. Thorndike, E. L.: The Relation between Memory for Words and Memory for Numbers, and the Relation between Memory over Short and Memory over Long Intervals. *American Journal of Psychology*, Vol. XXI, pp. 487–489.
11. Thorndike, E. L.: Memory for Paired Associates. *Psychological Review*, Vol. XV, pp. 122–138.
12. Whipple, G. M.: The Effect of Practice upon the Range of Visual Attention and of Visual Apprehension. *Journal Educational Psychology*, Vol. I, pp. 249–262.

TRUMAN L. KELLEY MEASURES MENTAL TRAITS[1]

C. S. SLOCOMBE

Boston Elevated Railway, Boston, Mass.

Statistical theories and methods applicable to results of mental testing have received a most noteworthy addition and extension by the publication of Kelley's latest book, "*Crossroads in the Mind of Man.*" Before proceeding to consider in detail its contents, it may be well to show its relationship to this previous work of which it constitutes an extension.

Prior Developments.—For a considerable number of years Spearman has been concerned with proving the existence of a general ability (*g*) underlying all abilities. According to his theory a person's ability in a particular test is due to a general ability or abilities specific to that particular test. Correlation then between a number of test results would, according to him, be due solely to the fact that the general ability enters into all of them. In the course of his work, he found, however, that as between any two or more tests, there might be some extra correlation not attributable to the *g* factor. He has assumed that some similarity in material or function has caused this added correlation. Insofar as this has interfered with his main purpose he has rejected the tests in which the added correlation showed such limited or group factors. More latterly, however, Spearman has considered the existence of his general factor proven and has gone on deliberately to the study of these group factors. His statistical techniques, however, have permitted him only to determine their existence and not their magnitude.

Kelley's Quantitative Analysis of Abilities.—Kelley's contribution in the volume under discussion is his presentation of the statistical procedures by which the presence and magnitude of such group factors may be determined, and any ability quantitatively analyzed completely. Chapter III is devoted to proving basic propositions and deriving formulæ for the determination of group factors or traits, as Kelley calls them. His propositions fifteen and sixteen contain the sections which will be most useful in determining contributing factors. The use of these formulæ is exemplified in Chapters IV, V, VI, and VII.

[1] Kelley, Truman L.: "Crossroads in the Mind of Man. A Study of Differentiable Mental Abilities." Stanford University, California: Stanford University Press, 1928. Pp.VII + 238.

For instance by means of them Kelley finds that the ability of a subject in Grade VII in tests of memory of meaningless symbols and speed of spatial relationships is due to the following factors in the following proportions:

	Memory of meaningless symbols	Speed of spatial relationships
(*a*) Maturity, heterogeneity, etc.	.385	.602
(*b*) Verbal		
(*c*) Arithmetic	.072	
(*d*) Memory	.480	−.232
(*e*) Spatial	.595	.457
(*f*) Speed	.084	−.287
(*g*) Vivacity	.123	−.188

But in Grade III, however, the factors determining ability in the same tests are somewhat different as shown below.

	Memory of meaningless symbols	Speed of spatial relationships
(*a*) Maturity, heterogeneity, etc.	.531	.319
(*b*) Verbal	.313	.195
(*c*) Arithmetic		
(*d*) Memory	....	−.208
(*e*) Spatial 1	.630	.153
(*f*) Speed		
(*g*) Spatial 2	....	.703

There is no necessity to discuss the procedures by which Kelley obtains these results, except to say that they are rather involved and necessitate a very considerable amount of tabulation and calculation and when arrived at are the results of long successive approximations. The possibilities opened up in the way of determining factors contributing to success in different tests might, were it not for certain considerations discussed below, be regarded as of the utmost importance. Whatever else may be said, the statistical techniques are undoubtedly of exceptional interest and value as extensions of theoretical psychology.

Magnitude of Errors Involved.—Karl Pearson[1] and E. B. Wilson[2] have pointed out in connection with Spearman's work the fact that some regression equations involve probable errors of such magnitude that the results obtained by these procedures are really indeterminable. For example, Spearman's equation for determining the amount of *g* and *s* contained in a score has not been found of practical value. Suppose by its use a subject's score in a particular test be divisible thus:

Score = general ability and specific ability.
80 = 60 + 20

The probable errors of the division are so large that the general ability factor may really be anywhere from 20 to 100 and the specific ability factor anywhere from −40 to +60! Results of this nature, are, of course, quite meaningless. Now as Kelley's procedures are even more elaborate, it is probable that his values which are quoted above are just as meaningless. For example, his final factors for Grade III are obtained at the close of the forty-fifth successive approximation with a population of only 60. The statistical technique presented by Kelley, then, is much too intricate to apply to data obtained from the crude and inaccurate methods of mental testing at present available. His work is not peculiar in this respect, and must by no means be regarded as of only academic interest for this reason. It is most probable that further study along similar lines will lead to simplification of procedure, for mathematical formulæ are almost always in their initial derivation clumsy and awkward. It is also to be hoped that the necessities of statistical treatment will inspire experimentalists to improve their methods.

The indeterminable nature of results at present obtainable by use of these formulæ must be kept in the foreground. Most people are easily "fooled by figures," and even Kelley though he explicitly states that his demonstrated applications of the formulæ are merely for purposes of exemplification rather than the determination of traits, enters into critical discussion on the basis of his findings.

Interpretation of Common Factor.—Kelley finds one common factor running throughout his tests. This corroborates the results of Spearman, that there is something common to all tests. However, Kelley refuses to accept Spearman's interpretation of this as a common factor

[1] *Biometrika*, December, 1927.
[2] *Science*, March 2, 1928.

having some physiological or psychological basis but considers it as due to heterogeneity in populations tested and seems to think that, if homogeneity were obtained such a common factor would not appear. He does not give much evidence bearing on this point but presents it merely as a difference in interpretation. It is, however, evident that Kelley's groups were heterogeneous to a degree which might have caused correlation, though one can see no particular reason why the degree should have varied so much in different tests. Kelley's explanation of these differences in terms of the effects of nurture on natural ability is not based on corroborative determinable evidence. On the other hand, a close examination of the work of Spearman indicates that homogeneity of groups was present to an extent that makes positive his proof of the existence of a psychological or physiological general factor. It may or may not be assumed that such factor entered into Kelley's results—but the heterogeneity of his groups prevents proof.

Homogeneity and Size of Groups.—This brings up the general question of homogeneity and size of groups. Generally it may be stated that the more homogeneous the group is required to be, the smaller must be its size, and *vice versa*. There does not seem to be any way of getting around this fact. Heterogeneity of a group may or may not be significant to the results or their interpretation, but always some assumptions have to be made. For example a group of children eleven to twelve years of age will include some from Grades IV, V, and VI. They are, then, heterogeneous in respect of scholastic training. An assumption that the amount of training which a child receives is determined by his ability to profit by training, and that the measurement of ability of pupils of uniform age differing in respect of training really measures the individual differences in their ability (the differences in training serving only to exaggerate or magnify these differences) gives apparent homogeneity. Similar assumptions, of course, have to be made when one grade, homogeneous in respect of age, is tested. As soon as groups from different schools, differing as to age, race, sex, etc., are combined, heterogeneity runs riot.

Alternative methods of dealing with this are, however, possible. A procedure which may be adopted in overcoming this difficulty (it has been used by Spearman) is to test separately a number of small groups as homogeneous as possible, involving as few assumptions as possible. As each group is homogeneous within itself but differs in some respect from every other group if results are corroborative,

proof may be regarded as satisfactory, where figures are determinate in the sense discussed above. Another method (used by Kelley) is to combine groups into one large group, admitting such heterogeneity as is unavoidable, and to interpret statistical findings with the knowledge of the heterogeneity. The first method gives results which admit of more exact interpretations but are incapable of use with elaborate statistical procedures. The second method on the other hand gives results which leave more room for diverse interpretations, but are capable of more involved treatment. The first method is not so reliable statistically, but the second method not so certain in the correctness of interpretation. Kelley's criticisms of published results rather look as if he has not clearly differentiated between these two inadequacies of method, or has required impossibly large, impossibly homogeneous groups before accepting results.

Conclusion.—Very many points raised by Kelley, and by examination of Kelley's work are worthy of discussion—for instance: the extent to which his factors of verbality, spatiality, etc., exist outside his or any other tests; whether his rigidity of requirement before accepting published results, though entirely right, is at all possible; whether the results, published in this book would meet the requirements of its author and be accepted by him; his statement that Spearman's Theory—as stated—is impregnable; etc., etc.

In fact almost every page in the book presents some material for consideration, leading either to discussion or agreement. This is quite in addition to the statistical presentations which are of the utmost interest. The main trouble is that the reader comes away with a feeling of hopeless wonder as to whether statistical psychology is not developing into a sterile, structural-functionalism.

Theoretical physics—an analogous and equally fantastic study—has produced results at the hands of experimentalists which have been of incalculable practical value, as for instance in industrial metallurgy. This may give hope that statisticopsychology may develop into something useful, but there is no guarantee, for psychology has been cluttered up with *objets d'art* ever since the days of the scholastics.

Spearman, and now Kelley, have made theoretical contributions leading to the analysis of abilities. Neither has done any more experimental work than is necessary to illustrate their methods. Until these methods are seriously applied by psychologists concerned with results, and their value proven thereby, they must be regarded as—just exceedingly interesting possibilities.

NOTES ON ARTICLES IN EDUCATIONAL PSYCHOLOGY IN CURRENT ISSUES OF OTHER MAGAZINES

REPORTED BY JAMES E. MENDENHALL
Research Associate, Institute of Educational Research, Teachers College
Columbia University

Intelligence Testing

General Review and Summary: Intelligence Tests. R. Pintner. Psychological Bulletin, July, 1928, 389–406.

Racial Differences in Intelligence in Minnesota. Margaret Lima. The Journal of Abnormal and Social Psychology, April-June, 1928, 68–71.

The Intelligence of Mexican School Children. Thomas R. Garth. School and Society, June 30, 1928, 791–794. A report of the National Intelligence Test administered to one thousand four Mexican children, Grades III to VIII. The average IQ of the group was 78.1; the IQ increases with school grade.

The Ten Year Level of Competency. Miles Murphy. The Psychological Clinic, May-June, 1928, 33–60. In the Grades, VA and VB, five hundred children were tested with the Witmer Formboard and Cylinders, Dearborn Formboard, in audito-vocal memory span forward, visual-vocal memory span forward, audito-vocal memory span reverse; the Stanford Revision of the Binet-Simon Intelligence Scale; and an arithmetic test.

A Study of Orphan Children in Texas. The Journal of Delinquency, March, 1928. One hundred fifty-one orphanage children ages six to twenty-one, secured an average IQ of 85.5 (Dearborn) and 86.2 (Haggerty); five hundred four public school children, an average IQ of 97 (Dearborn) and 98.7 (Haggerty).

The Organization of Mental and Physical Traits during Adolescence. Fowler D. Brooks. The Journal of Applied Psychology, April, 1928, 228–229. Data derived from physical and mental measurements of one thousand one hundred eighteen individuals of chronological ages from thirteen to twenty years are presented. In practically all cases correlations between physical and mental ages (*CA* constant) are zero.

The Relation of the Intelligence of Pre-school Children to the Occupation of Their Fathers. Florence L. Goodenough. The American Journal of Psychology, April, 1928, 284–294. "A group of three hundred eighty children, all of whom were between the ages of eighteen and fifty-four months, were given the Kuhlman Revision of the Binet tests twice, the two examinations being separated by an average interval of six weeks." The Barr Scale for occupational intelligence and the Taussig Industrial Classifications were used.

A Study of the Relation of Intelligence Test Scores of Public School Children to the Economic Status of Their Parents. J. B. Stroud. Journal of Genetic Psychology, March, 1928, 105–111. Correlations between test scores and tax assessments (*N*38 to 106) range from +.10 to +.35.

How Athletes and Non-athletes Compare in Mental Ability and in Educational Achievements. Robert T. Hall. American Physical Education Review, June, 1928, 388–398. In the study one hundred eight boys and seventy-six girls (high School) were tested by the Terman Group and the Iowa High School Content Examination. Comparisons were made between athletes and non-athletes in age, in extra-curricular participation, in intelligence and in achievement. Boy athletes are inferior to non-athletes in intelligence and achievement, but are superior in chronological age and in extra curricular activities. The converse is true with girls by similar comparisons.

Mental Test Traits of College Cribbers. Herbert C. Brownell. School and Society, June 23, 1928, 764. The writer concludes regarding thirty cribbers in final examinations that: (*a*) fifty-eight per cent are below the campus average in intelligence, (*b*) seventy-one per cent are more extroverted than the campus average, and (*c*) eighty per cent are more psychoneurotic than the campus average.

IQ and Ability in Art. Alfred L. Lewerenz. School and Society, April, 21, 1928, 489–490. The Los Angeles Tests in Fundamental Abilities of Visual Art and a group intelligence test yield correlations ranging from .009 to .275 (*N*674 to 995).

Norms for the 1927 *Psychological Examination.* L. L. Thurstone. The Educational Record, April, 1928, 102–107.

Measurement of Achievement

General Review and Summary: Educational Tests. Vernon Jones. Psychological Bulletin, July, 1928, 407–421.

The Measurement of the Effect of Teaching. S. A. Courtis. School and Society, (I) July 14, 1928, 52–56; (II) July 21, 1928, 84–88. The writer presents a method which attempts to determine the effect of teaching apart from maturation. The method was applied to an experimental situation involving the learning to spell words.

A Study of the Relationship of Scholastic Maturity and Scholarship as Shown by College Marks. Edward B. Mersereau. Peabody Journal of Education, March, 1928, 291–300. College marks increase as the amount of college work that a student has had increases.

The Reliability and Validity of Certain Arithmetic Tests. Peabody Journal of Education, May, 1928, 339–341. Coefficients of reliability are reported (*N* = 73) Woody-McCall .58 to .68, Pittsburgh Arithmetic .68 to .71, Buckingham Scale (Probs.) .56, Monroe Reasoning .45 and .55, Otis Arithmetic .75, Courtis .51, Cleveland .79, Stanford Achievement Adv. .87.

The Relative Accuracy of Ratings Assigned With and Without Use of Drawing Scales. Fowler D. Brooks. School and Society, April 28, 1928, 518–520. Fifteen specimens were rated by twenty-one judges using (*a*) no scale (*b*) the Thorndike Scale (*c*) the Kline-Carey scale. The coefficient of variability of the ratings indicated that the Thorndike scale yielded the most reliable ratings, the Kline-Carey yielding the next. "These data indicate that the use of a drawing scale decreases the inaccuracy of ratings to approximately one-half of what it is when no scale is used."

A Scale for Measuring Achievement in General Science. August Dvorak. School Science and Mathematics, June, 1928, 608–613. Reliability coefficients ranging from .70 to .877 are reported.

Some Factors of Success in Physics. Virgil C. Lohn. School Science and Mathematics, April, 1928, 389–398. Sixty high school and sixty-four college students were tested with a mathematics, an Otis and the Ruch-Poponoe tests, Intercorrelations between these, a reading test, and marks in the physics courses were obtained.

Prevalent Errors in New Type Examinations. D. A. Worcester. Journal of Educational Research, June, 1928, 48–52.

The Place of New Type Tests in Teaching Mathematics. William David Reeve. Teachers College Record, May, 1928, 693–703. A criticism of the old type and of the new types of tests in the field of mathematics.

Psychology of Learning and School Subjects

General Review and Summary: Studies in the Psychology of Children's Drawings. Florence L. Goodenough. Psychological Bulletin, May, 1928, 272–283.

An Experimental Investigation into Transfer of Training in Skilled Performances. J. N. Langdon and Edna M. Yates. The British Journal of Psychology, April, 1928, 422–437. "The trained (two weeks) group contained thirty-two and the control group twenty-eight subjects . . . The results of both experiments lead to the conclusion that in manual dexterity and in visual discrimination, the effect of training is specific rather than general. In some cases, the brief practice afforded by the test itself is shown to be more effective than the intensive training in an analogous performance."

Finding Mistakes vs. Correct Associations in Simple Number-learning. Garry C. Myers and Caroline E. Myers. Journal of Educational Research, June, 1928, 25–31. "Five hundred sixty-four elementary school children were tested individually." The correct-association method proved definitely superior both in time for learning and in correctness of responses on test.

The Rôle of Articulation in Memorizing. M. C. Barlow. Journal of Experimental Psychology, August, 1928, 306–312. Three hundred twenty-seven cases in Grades II to IX and in college were practiced and tested in memory for non-sense syllables. Verbalization proved superior for each grade group, the superiority being greater for the higher grades.

Verbal Memory in the Pre-school Child. Josephine C. Foster. Journal of Genetic Psychology, March, 1928, 26–44. Thirty-one children age 2–7 to 4–9 composed the experimental group. The children of the older mental ages reproduced a greater amount of the material.

Summary of Arithmetic Investigations. Guy Thomas Buswell. The Elementary School Journal, May, 1928, 702–709. June, 1928, 703–742.

When Should We Teach Arithmetic? A Committee of Seven Investigations. Carleton Washburne. The Elementary School Journal, May, 1928, 659–665. In one hundred twenty-five school systems wide variations were found in the grade in which a topic was introduced and completed . . . Sixth grade pupils who first had formal arithmetic in the first grade were superior to those who first had arithmetic in the second or in the third grades.

Are Any Number Combinations Inherently Difficult? Carleton Washburne and Mabel Vogel. Journal of Educational Research, April, 1928, 235–255. Three forms of an arithmetic test (addition, subtraction, multiplication, division) were given to eight hundred forty-five Winnetka children. Form I easiest to hardest; Form II hardest to easiest; Form III hardest, easiest, next hardest, etc. The performances are compared with those of Wisconsin children studied by Clapp.

Analysis of Errors in Fractions. Leo J. Brueckner. The Elementary School Journal, June, 1928, 760–770. The 21,065 errors were classified into a few major causes.

A Revised List of Phonics for Grade II. Carleton Washburne and Mabel Vogel. The Elementary School Journal, June, 1928, 771–777. Arthur I. Gates, "A Reading Vocabulary for the Primary Grades" and a study of phonics by the authors in 1923 are used as bases for a revised list of phonics.

The Newark Phonics Experiment. Elmer K. Sexton and John L. Herron. The Elementary School Journal, May, 1928, 690–701. "Nearly one thousand pupils were included in the experiment . . . The teaching of phonics functions very little or not at all with beginners in reading during the first five months. It begins to be of some value during the second five months but is of greater value in the second grade."

The Effectiveness of Drill in Handwriting to Remove Specific Illegibilities. Hilda Lehman and Luella C. Pressey. School and Society, May 5, 1928, 546–548. The handwriting of the experimental group of twenty-three, Grade IVB children is compared with that of nineteen, Grade IIIA children. Remedial drill regarding specific illegibilities increased the rate, quality, and legibility significantly of the writing of the experimental group in comparison with itself before practice and with the control or general practice group.

An Association Study of the Vocabulary of Grade Children. C. G. Shambaugh and Olive L. Shambaugh. Journal of Educational Research, June, 1928, 40–47. One thousand fifty-one children, Grades IV to VIII, gave a total of 230,631 words in a controlled association test. A vocabulary based on the most frequent responses was thereby constructed.

Difficult Words and the Beginner. Will E. Wiley. Journal of Educational Research, April, 1928, 278–289. Fifty-six first grade children were practiced and tested for a period of twelve weeks. The results tended to show little correlation between difficulty of learning a word and its rating by either the Thorndike or the Gates word lists.

A Vocabulary Study of Children in a Foreign Industrial Community. Alice M. Jones. The Psychology Clinic, March, 1928, 13–21. A group of five hundred fifty-two school children up to the Grade VIII over eight years of age were administered the first list of words in the Stanford Revision.

Numerals vs. Words for Efficiency in Reading. Miles A. Tinker. The Journal of Applied Psychology, April, 1928, 190–199. An investigation of the eye movements of forty college students showed "that the compactness of printing achieved by employing numerals rather than words brings about faster reading when other conditions remain approximately constant."

An Analytical Study of Basal Reading Texts for the Elementary School. Mabel E. Simpson. The Journal of Educational Method, June, 1928, 400–411. An outline of a method for examining texts. A score card is presented.

Character and Personality Traits

General Review and Summary: Personality and Character Traits. Mark A. May, Hugh Hartshorne, and Ruth Welty. Psychological Bulletin, July, 1928, 422–443.

A Test for Ascendance-submission. Gordon W. Allport. The Journal of Abnormal and Social Psychology, July-Sept., 1928, 118–136. The test reactions of five hundred college students to certain described social situations is validated by combined self and fellow ratings on an ascendance-submission scale (r = .29 to .633). Reliability coefficients of test. 737 to .78.

Ascendance-submission in Relation to Certain Other Factors in Personality. Irving Edison Bender. The Journal of Abnormal and Social Psychology, July-Sept., 1928, 137–143. On four hundred college students the following correlations are secured between ascendance-submission scores and (*a*) heights, −.03 (*b*) weights, +.09 (*c*) intelligence, +.0008 (*d*) scholarship, +.062 (*e*) extroversion (Heidbreder Test), +.379. "No reliable differences in mean scores appear as a result of intelligence, scholarship, position in the family, or class in the college."

A Biochemical Approach to the Study of Personality. Gilbert J. Rich. The Journal of Abnormal and Social Psychology, July-Sept., 1928, 158–175. "The correlations between chemical determinations and personality traits obtained in this study (of fifty-seven college students and three hundred three children) have practically all been low, between .20 and .30."

A Pre-school Character Rating Chart. K. M. Banham Bridges. The Psychological Clinic, May-June, 1928, 61–72. On a scale of fifty items, twenty nursery school children are rated. Age differences were found. An average rank correlation between ratings of different raters on the group was +.78.

Stealing as a Form of Aggressive Behavior. E. M. Riddle. The Journal of Abnormal and Social Psychology, April–June, 1928. A study of the mental age, the frequency, the kind, the place of stealing, behavior deviations, and character traits of one hundred ninety children (ages 4-0 to 22-0) who had committed petty thefts.

Statistical Methods

The Scoring Reliability of Test Material in the Free Answer and Short Answer Forms. Public Personnel Studies, May, 1928, 98–108.

Suggested Tests to Measure Knowledge of Statistical Principles and Procedures. Public Personnel Studies, May, 1928, 109–115. The test covers these aspects of statistics: information, terms, and practices, judgment, situations, calculations, understanding directions, and memory for oral directions.

The Absolute zero in Intelligence Measurement. L. L. Thurston. The Psychological Review, May, 1928, 175–197.

The Theory of Factors. Stuart C. Dodd. May, 1928, 211–234. Part II, July, 1928, 261–279. A critical discussion of "all the original contributions to Spearman's theory of two factors, '*g*' and '*s*,' and to Thomson's sampling theory."

Exceptional Children

Results of Examination of Ninety-one Girls at a State Institution. The Psychological Clinic, March, 1928, 22–28. "The results show marked pedagogical retardation; but a normal display of intelligence."

The Influence of the Group upon the Stutterer's Speed in Free Association. Lee Edward Travis. The Journal of Abnormal and Social Psychology, April-June, 1928, 45–51. Eight out of ten stutteres did better alone than with the group, fourteen out of fifteen normal speakers did better with the group than alone.

The Problem of Childhood. L. Daniel House. The Journal of Delinquency, March, 1928, 25–47. A questionnaire (eighty items, thirty items revised) regarding childhood behavior was submitted to one hundred twenty psychoneurotics and five hundred forty normals. The normals reported a greater number of "symptoms" during childhood (such as "fidgeting," day dreaming, desire to run away from home) than did the psychoneurotics.

A Survey of Problem Pupils. Gertrude Hildreth. Journal of Educational Research, June, 1928, 1–14. An attempt to isolate certain traits possessed by so-called "problems" pupils.

Clinical Studies in Reading. H. L. Donovan. Peabody Journal of Education, March, 1928, 269–277. Suggested methods for remedial work in reading.

NEW PUBLICATIONS IN EDUCATIONAL PSYCHOLOGY AND RELATED FIELDS OF EDUCATION

CONDUCTED BY FRANCES M. FOSTER

LOOKE IN THY HEART

Foole, said my Muse to me, looke in thy heart, and write.

—PHILIP SIDNEY.

The New Leaven, by Stanwood Cobb. New York: The John Day Co., 1928. Pp. 340.

In *The New Leaven*, Stanwood Cobb has written an important, and what is even more important, an extremely readable interpretation of the whole progressive movement in education. His manner is warm and conversational rather than lectury. He has his facts straight. And if he is biased, it is so decidedly in favor of the human as against the mechanical in teaching and administration, that he needs no apologist. In brief, he pleads Jesus rather than Hegel: Schools made for men, and not men for schools. The book holds an infinity of interests, ranging all the way from challenge to comfort, for parents, teachers, administrators who hold the power of help or hindrance over youngsters going over the road from the nursery to the university.

There are many excellent things in *The New Leaven*. The author sees soundly enough that education is not limited to the school-room alone; that the home is a vital and powerful factor. He gives due credit to modern parents for making progressive education possible. In his own words, "The trend of education in any age is after all what parents desire it to be. It is the expression of their intelligent vision or lack of vision." In a rapid, vivid review of the inception of the movement in this country, he pays tribute to the pioneers, Francis W. Parker, John Dewey, Mrs. Emmons Blaine, Marietta L. Johnson and others, who did so much more than the spade work in the period from 1894 to 1917. Parker's contribution he sums up as, "the school . . . to be adjusted to the child rather than the child to the school." Dewey's he sees as two-fold, "First, he held that . . . activity programs in the education of children are of the greatest importance; and secondly, he emphasized the value, with children, of the exchange of ideas and experiences in a social medium."

Sensibly enough, Cobb makes no attempt at a definition of "progressive" education. Instead, he formulates a sort of decalog of principles based on descriptions offered by thirty-two progressive teachers from schools all over the country.

(1) Health Must Come First. (2) Learning Comes from Doing: Let the Hand Aid the Brain. (3) The Classroom Should Be Freed from Unnatural Restraints. (4) Adapt Education to the Differences of the Individual Child. (5) Group-consciousness and Social-mindedness Should Be Developed in Children. (6) The Child Should Have Abundant Opportunity for Creative Expression. (7) Enable the Child to Acquire Control of the Tools of Learning Rather than Merely to Acquire Facts. (8) Introduce into Academic Work the Method of Creative Expression, so that Education Shall Be Joyous. (9) Abolish the Tyranny of Marks and Examinations. (10) The Teacher Should Be a Leader and Guide, Not a Taskmaster.

With these ten principles as text, he writes ten brilliant, indignant, amusing chapters, full of excellent human stuff of what most schools are, what all schools might be and ought to be, and some few, very few schools are decently approximating toward the progressive ideal. In those ten chapters a reviewer is faced with an embarrassment of riches and would do a poor service to the prospective reader if he went into detail. Oh, excellent! But for the last three chapters comment in detail is almost imperative. For these seem really the nub of the book. Here Cobb takes a fall out of the colleges, and lays at their door most of the current evils in present day education. He sees them as succubi, draining the life out of all that is vital in teaching in America. Himself a quondam examiner, he sees the College Entrance Examinations as a smart-aleck guessing game, in which one wins or loses by rule, in tests that test only the candidate's memory and the examiner's ingenuity. Free the secondary school from the tyranny of the college, is his cry. Liberalize the examinations or, what is even more desirable, go ahead after the real, the living in education, and if the colleges are not suited, so much the worse for them. The real solution, however, to Cobb, lies in the Junior College; not as a mere link between secondary school and college proper, but as a whole, rounded, adequate adolescent education in itself, and therefore, *ipso facto*," a distributing center for adult education in each locality."

B. J. R. Stolper.

The Lincoln School, Teachers College.

HANDBOOKS FOR THE RESEARCH WORKER IN EDUCATION

How to Research in Education, by Carter V. Good. Baltimore: Warwick and York, Inc., 1928. Pp. XII + 298.

Check Lists for Educational Research, by Harold Hench Bixler. New York: Bureau of Publications, Teachers College, Columbia University, 1928.

It is common knowledge that during recent years educational research has increased tremendously both in amount and in complexity. Often each worker in the field used his own peculiar methods of organizing and presenting the problem which he had investigated. The novice in the field of education was "at sea"—he lacked sources to which he might turn to find suggestions as to methods and tools.

In "How to Do Research Education," the student will find many helpful chapters: Sources of information in education; technical vocabulary and terminology; characteristics of scientific investigations; selecting and defining a problem for investigation; techniques for the collection of data; problems of analysis, organization, and summarization of data; interpretation of data and formulation of conclusions. This book is especially well adapted to orienting the student. Its author has canvassed magazines, monographs, and books in education in a most comprehensive manner; the important points discovered by this survey have been presented in a concise and readable style.

"Check Lists for Educational Research" is another book which every student interested in research should find useful. The writer has systematically analyzed and compiled the comments, the statistical devices, etc. which have appeared in scientific educational literature during the last five years. These points have been organized into five chapters: Selection and formulation of a research problem; methods of research; techniques used in research; statistical methods used in research; the presentation of the research.

The research worker will find that this list of key points will aid him in avoiding pitfalls of procedure. Moreover, he will discover here many valuable suggestions bearing on problems of method, of testing, of statistics—devices which he might otherwise fail to consider.

Indeed, these two books well complement each other. The one introduces the reader to the fields and methods of research; the other furnishes a definite and ready reference for concrete suggestions relative to the solution of a given problem. For furthering the important function of research, these books should be of signal service to every advanced student of education.

JAMES E. MENDENHALL.

The Lincoln School, Teachers College.

INFANCY

Infancy and Human Growth, by Arnold Gesell. New York: The Macmillan Company, 1928. Pp. XVII + 418.

Extending the methods of study used in the earlier volume, "Mental Growth of the Pre-school Child" (1925), Dr. Gesell deals at closer range with the characteristics of early human growth as determined by systematic observation and direct record. The studies upon which this volume was based were made in the Yale Psycho-clinic under the direction of Dr. Gesell.

Following an introductory chapter which deals in a general way with the total cycle of mental growth, the book is divided into three parts. Part I sketches the methods and technique of observation of infant development. Through periodic and systematic application of these observational methods, monthly increments of development are determined from which a tentative infant-development recording schedule is drawn. This is in the form of a normative summary giving the important monthly and tri-monthly increments of growth. Here in sixteen progressive developmental stages, between the ages of one month and three years, are presented "a systematic series of typical behavior pictures along the advancing cycle of growth." Approximately a dozen items covering motor, language, adaptive, and personal-social aspects of behavior are recorded. The scale records only normal mental growth and does not deal with subnormal or abnormal growth phases. It should prove valuable in all fields where it is necessary to follow clinically the mental growth of the human young.

Part II treats of genetic studies of infant behavior. Developmental studies of over one hundred infants have been secured in order to determine the individual trend and tempo of development. Here are discussed, with case study examples, normal and retarded development, the symptoms of giftedness in infancy; atypical and pseudo atypical growth complexes; drawing (crayon-paper behavior) as a developmental index; glandular and nutritional factors in mental growth such as mongolism, puberty præcox, thyroid deficiency, and rickets. The developmental aspects of the physiology of twinning, and the consequences of premature growth are also given a place in the discussion.

Dr. Gesell believes firmly in the close relation between psychological and biological aspects of developmental problems, but he also believes that growth in itself is so vital a factor that its mental phases need not be permanently impaired in every case of motor disability. Citing

case studies of infants who, in spite of motor disabilities due to birth injury, still showed an ascendancy in mental growth, he says: "It appears that the organization of the striped musculature is not as extremely essential to the attainment of mental maturity as has been suggested by over-mechanical theories of human behavior."

Part III is devoted to a discussion of the significance of infancy. Problems of heredity in relation to mental growth are discussed. The interdependence of heredity and environment are brought out, the early genesis of individual differences is emphasized and there is postulated "a progressive kind of determinism which, in the field of behavior, at least, comes under human control, and is inconsistent with a fatalistic view of infancy." The possibilities and limitations of developmental predictions in the clinical study of mental growth are considered. The book concludes with a chapter describing growth as a medical and social concept.

The book is excellent as to type, set up, and illustrations, making a handsome as well as useful volume. It is a real contribution to a field of investigation still quite new to science.

ANN SHUMAKER.
Lincoln School, Teachers College, N. Y.

A TEACHER'S PROBLEM BOOK IN EDUCATION

A Word Book Syllabus in Principles of Education, by Orlie M. Clem. Baltimore: Warwick and York, 1928. Pp. 170.

This book is a pioneer effort to apply progressive principles of teaching to work on the college level. It is designed to get students to expend a maximum of energy in thinking about general principles of education. It can be used as a syllabus in classes in "Principles of Education" or similar summarizing courses in the junior or senior year in teachers colleges or schools of education.

A more explicit title would have been "Work Book Syllabus in Principles of *Secondary* Education," since many of the problems and units pertain directly to high school work. For example, there are units on the "Secondary School Pupil," and on "Adolescence." The author himself says in his foreword that, "The major emphasis is given to the field of secondary education."

The book is introduced by an essay entitled "A Philosophy of Teaching." The essay presents the philosophy underlying the "Work Book." The main point in this philosophy is that a teacher's function "is to present challenges to students and to help them to

move away from the crossroads" which these challenges invite them to face. The tenor of the essay gives the impression that there is a fear that the main body of the book needs a defense and would be unable to stand on its own merits.

The main body of the book is divided into units of work under twenty titles such as "Individual differences," "Aims of Education," "Guidance," "The Teacher," "The New Education," etc. Each unit consists of a series of questions and problems related to the topic of the unit and is accompanied by a selected bibliography. Many of the problems are excellent. The practice obtained through thinking about and discussing these problems should enable students to work out a stable set of principles by which to guide their own teaching.

Two of the units do not fit in well with a book dealing with principles of education. Unit II, "The physiological basis of learning," goes into too much detail and relates to material which should be learned in psychology. Unit XX, "Leading educators," asks: "What has been the specific contribution of each of the following men to education? Where are they now located?" These questions place emphasis on memorizing facts of subordinate importance and only incidentally bring in principles of education.

In selecting these units of study, perhaps more emphasis should have been placed on the relation between education and the public in such matters as the support and control of schools, and less on the psychological problems involved in teaching. However, "Principles of Education" is a broad field and it is to be expected that there will be wide differences of opinion regarding the points most deserving of emphasis.

It is unfortunate that a book of such promise should be somewhat lacking in careful organization and writing.

Nevertheless, it is a very worthy effort to put into practice those principles which all too often instructors in education courses only preach.

NEAL BILLINGS.
State Teachers College, Milwaukee, Wisconsin.

A TEACHER'S BOOK IN MENTAL HYGIENE

The Hygiene of Instruction, by Lawrence Augustus Averill. Boston: Houghton Mifflin Company, 1928. Pp. XIV + 386.

"A Study of the Mental Health of the School Child" is the subtitle of this book. In a brief introduction by Dr. Cubberley, he recom-

mends it as a textbook in mental hygiene for teacher-training institutions.

The author, who is convinced that the "prevention of distorted and disharmonious personalities is achievable," attempts to cover all aspects of the many problems of mental hygiene of the school child. Nor is this all. He also includes a treatment of all aspects of the school situation which have mental hygiene significance.

Some notion of the vastness and inclusiveness of the undertaking can be obtained from the chapter captions. The first seven chapters treat of the more general aspects of the problem and are evidently intended to serve as a psychological background. They are captioned as follows: The Hygiene of Attitude; The Hygiene of Performance; The Mental Hygiene of the School Day; The Physical Basis of Mental Health; Habit and the Conditioned Reflex; and two chapters on The Hygiene of Classification—(1) the mentally deficient child, and (2) the gifted child. The next two chapters treat of varieties of problem children, including: (1) The deceitful, (2) the disobedient, (3) the emotionally abnormal, (4) the indifferent, (5) the irresponsible, (6) the lazy, (7) the lacking in self-control, and (8) the selfish. Chapters X and XI are captioned Mental Hygiene and the School Subjects. Study, Home Sources of Conflict, Mental Hygiene, The New Education and the New Teacher, and The Child-guidance Clinic, respectively, are the subjects treated in the remaining chapters.

In a book which covers so many aspects of a field of knowledge which is itself an integration of many fields, it is, of course, impossible to treat any one aspect exhaustively. In his treatment of such subjects as surveys of delinquents, character education and the establishment of child guidance clinics, no mention is made of some of the most important studies or activities related to these fields. His treatment of the mental hygiene of school subjects is disappointingly inadequate. Intentionally or not, it neglects the consideration of the adaptability of materials of instruction to individual needs and capacities. And from the point of view of mental hygiene the need is to integrate the traits the child has, not emotionalize over the great possibilities of integrating the most desirable ones without regard to the child's possessions. But considering the fact that the book is purposed to "aid teachers and inspire parents," that, to serve the latter, a bowing acquaintance with knowledge will suffice,—and that it will be used largely in teacher-training institutions—a more technical, scholarly and comprehensive treatment is perhaps inadvisable.

If it was the wish of the author to inspire by what Kidd calls the "contagion of emotions" he has probably succeeded. There is no lack of high sounding statements in this book. The manner is persuasive—though at times somewhat forced, overdone, much too wordy. But with all its omissions and faults it is to date the most inclusive treatment of the subject. A reading of this volume is bound to give a student an insight into the possibilities of treating peculiarities of behavior in children which he is not likely to get in other courses in education. It should prove especially useful in institutions which are not adequately equipped with books and journals which treat of mental hygiene problems.

H. MELTZER.

Psychiatric Clinic, City of St. Louis, Mo.

COLLEGE TEXTBOOK IN PSYCHOLOGY

General Introduction to Psychology, by C. R. Griffith. New York: The Macmillan Company, 1928. Pp. 591.

The aims of this book are (*a*) to give students a broad prospective over the whole field of psychology which they are assumed to have known before in terms of fundamentals and (*b*) to give them a reasonable orientation in the field of psychology at large. Very seldom the contents of textbooks in psychology have any relation to the aims stated in the preface. This book constitutes a noteworthy exception. In that it does, it is essentially a college textbook and makes no pretense at being written for both college students and the public.

It discusses the so-called foundations of psychology; gives a historical sketch of experimental psychology and some discussion of principles; it touches on animal psychology, intelligence, social psychology, abnormal psychology, and psycho-technology. The discussion of these various aspects of psychology appears inclined at times to be a little scrappy. In such an encyclopediac text this is more or less inevitable. It also is probably due to the fact that references quoted are only used for illustrative purposes. The illustrations for the most part are of points of view and are well chosen.

The outstanding value of the book is the exceedingly reasonable point of view of the author. That is to say, he does not line himself up as a behaviorist, Freudian, or what not, but discusses rather, in an impartial manner, the logical truths and fallacies contained in the

various doctrines. This is liable to give the student an understanding of the fact that psychology is in the painful process of evolving toward a rational science. Insofar as the tone of the book implies optimism that psychology will evolve in the right direction, the student, if he adopts its point of view, will be encouraged to continue in the study and, in his continuance, avoid bigoted pedantry.

CHARLES S. SLOCOMBE.

Boston Elevated Railway, Boston, Mass.

PUBLICATIONS RECEIVED TO SEPTEMBER 15, 1928

MONOGRAPHS

SNYDER, AGNES: *The Value of Certain Measurements in the Training of Teachers.* Baltimore: The Johns Hopkins Press, 1928, pp. IX + 146.

DOUGHERTY, MARY LORETTE: *A Comparative Study of Nine Group Tests of Intelligence for Primary Grades.* Baltimore: The Johns Hopkins Press, 1928, pp. VII + 111.

THOMAS, MINOR WINE: *Public School Plumbing Equipment.* New York City: Bureau of Publications, Teachers College, Columbia University, 1928, pp. 128.

TIDWELL, C. J.: *State Control of Textbooks.* New York City: Bureau of Publications, Teachers College, Columbia University, 1928, pp. 78.

IRVIN, OSCAR WILLIAM: *State Budget Control of State Institutions of Higher Education.* New York City: Bureau of Publications, Teachers College, Columbia University, 1928, pp. VI + 122.

HOLLINGSHEAD, ARTHUR D.: *An Evaluation of the Use of Certain Educational and Mental Measurements for Purposes of Classification.* New York City: Bureau of Publications, Teachers College, Columbia University, 1928, pp. IX + 63.

SCHAAF, WILLIAM L.: *A Course for Teachers of Junior High School Mathematics.* New York City: Bureau of Publications, Teachers College, Columbia University, 1928, pp. VI + 160.

STURTEVANT, SARAH M. and STRANG, RUTH: *A Personnel Study of the Deans of Women in Teachers Colleges and Normal Schools.* New

York City: Bureau of Publications, Teachers College, Columbia University, 1928, pp. 95.

MEADER, EMMA GRANT: *Teaching of Speech in the Elementary School.* New York City: Bureau of Publications, Teachers College, Columbia University, 1928, pp. VI + 129.

HUBER, MIRIAM BLANTON: *The Influence of Intelligence upon Children's Reading Interests.* New York City: Bureau of Publications, Teachers College, Columbia University, 1928, pp. 39.

GRIFFIN, ORWIN BRADFORD: *The Evolution of the Connecticut State School System.* New York City: Bureau of Publications, Teachers College, Columbia University, 1928, pp. 261.

ZIEGLER, CARL WILLIAM: *School Attendance as a Factor in School Progress.* New York City: Bureau of Publications, Teachers College, Columbia University, 1928, pp. VI + 63.

ELSBREE, WILLIARD S.: *Teacher Turnover in Cities and Villages of New York State.* New York City: Bureau of Publications, Teachers College, Columbia University, 1928, pp. 88.

VAN DE VOORT, ALICE MARIA: *The Teaching of Science in Normal Schools and Teachers Colleges.* New York City: Bureau of Publications, Teachers College, Columbia University, 1928, pp. IV + 96.

UHRBROCK, RICHARD S.: *An Analysis of the Downey Will-Temperament Tests.* New York City: Bureau of Publications, Teachers College, Columbia University, 1928, pp. 78.

KNIGHT, FREDERIC B. and BEHRENS, MINNIE S.: *The Learning of the One Hundred Addition Combinations and the One Hundred Subtraction Combinations.* New York City: Longmans, Green and Compay, 1928, pp. XIII + 82.

SOCIAL, GENERAL AND APPLIED PSYCHOLOGY

TROLAND, LEONARD T.: *Human Motivation.* New York: D. Van Nostrand Company, Inc., 1928, pp. XIV + 521.

ELLIS, ROBERT S.: *The Psychology of Individual Differences.* New York: D. Appleton and Company, 1928, pp. XXIII + 533.

HUNTER, WALTER S.: *Human Behavior.* Chicago: The University of Chicago, 1928, pp. X + 355.

GATES, ARTHUR I.: *Elementary Psychology.* New York: The Macmillan Company, 1928, pp. XVII + 612.

ARLITT, ADA HART: *Psychology of Infancy and Early Childhood.* New York: McGraw-Hill Book Company, Inc., 1928, pp. IX + 228.

SYMONDS, PERCIVAL M.: *The Nature of Conduct.* New York: The Macmillan Company, 1928, pp. XI + 346.

FLOWER, J. CYRIL: *The Psychology of Religion.* New York: Harcourt, Brace and Company, Inc., 1928, pp. XI + 248.

RABAUD, ETIENNE: *How Animals Find Their Way About.* New York: Harcourt, Brace and Company, 1928, pp. IX + 142.

PIERCE, FREDERICK: *Understanding Our Children.* New York: E. P. Dutton and Company, 1928, pp. X + 198.

GENERAL EDUCATION

BURGESS, MAY AYRES: *Nurses, Patients and Pocketbooks.* New York: Committee on the Grading of Nursing Schools, 1928, pp. XII + 618.

Canadian Committee on Modern Languages: Modern Language Instruction in Canada. Volume I. Toronto: The University of Toronto Press, 1928, pp. XLVIII + 547.

Canadian Committee on Modern Languages: Modern Language Instruction in Canada. Volume II. Toronto: The University of Toronto Press, 1928, pp. V + 851.

WHEELER, WILLIAM MORTON: *The Social Insects.* New York: Harcourt, Brace and Company, 1928, pp. XVIII + 378.

BOVET, PIERRE: *De L'Efficacité De L'Exemple.* Paris: Librairie Philosophique J. Vrin, 1928, pp. 251.

SEARS, JESSE B.: *Classroom Organization and Control.* Boston, Mass.: Houghton Mifflin Company, The Riverside Press Cambridge, 1928, pp. XII + 400.

HARRIS, PICKENS E.: *Changing Conceptions of School Discipline.* New York: The Macmillan Company, 1928, pp. XVIII + 384.

MACKANESS, GEORGE: *Inspirational Teaching.* New York: E. P. Dutton and Company, 1928, pp. XI + 225.

GODDARD, HENRY HERBERT: *School Training of Gifted Children.* Yonkers-on-Hudson, New York: World Book Company, 1928, pp. X + 226.

HUDELSON, EARL: *Problems of College Education.* Minneapolis: The University of Minnesota Press, 1928, pp. XVI + 449.

GRIZZELL, EMIT DUNCAN: *Education: Principles and Practices.* New York: The Macmillan Company, 1928, pp. XVI + 428.

Elementary School Textbooks

Harper, Wilhelmina and Hamilton, Aymer Jay: *Pleasant Pathways.* New York: The Macmillian Company, 1928, pp. IX + 262.

Breed, Frederick, S. and French, William C.: *The Breed-French Speller.* New York: Lyons and Carnahan, 1928, pp. XIV + 257.

Hepner, Walter R. and Hepner, Frances K.: *Junior Citizens in Action.* New York: Houghton Mifflin Company, The Riverside Press Cambridge, 1928, pp. XII + 253.

Pribble, Evalin: *Correct English Usage.* New York: Lyons and Carnahan, 1928, pp. 255.

Lee, Ettie: *Living English Studies.* New York: The Macmillan Company, 1928, pp. XII + 319.

Washburne, Carleton: *Washburne Individual Arithmetic, Books One to Five.* Yonkers-on-Hudson, New York: World Book Company, 1927,

Washburne, Carleton: *Washburne Individual Arithmetic, Key for Text Book: Books One to Five.* Yonkers-on-Hudson, New York: World Book Company,

Books and Monographs on the Curriculum

Monroe, Walter S. and Herriott, M. E.: *Reconstruction of the Secondary-school Curriculum: Its Meaning and Trends.* Illinois, The University of Illinois, Urbana, 1928, pp. 120.

Buchanan, M. A. and MacPhee, E. D.: *An Annotated Bibliography of Modern Language Methodology.* Toronto: The University of Toronto Press, 1928, pp. 428.

Winslow, Leon Loyal: *Organization and Teaching of Art.* Baltimore: Warwick and York, Inc., 1928, pp. 243.

Maroney, Frederick William: *Physical Education for Public Schools.* New York: Lyons and Carnahan, 1928, pp. VIII + 397.

Teachers' Outlines and Helps for the Teachers

Harper, Wilhelmina and Hamilton, Aymer Jay: *Helps for the Teacher.* New York: The Macmillan Company, 1928, pp. VII + 24.

Raleigh Elementary Education Council: A Suggested List of Activities. Raleigh, N. Carolina: Bynum Printing Company, 1928, pp. XV + 66.

Tests and Measurements

Greene, Harry A.: *Work-book in Educational Measurements.* New York: Longmans, Green and Company, 1928, pp. VI + 156.

Baker, Harry J.: *Detroit Advanced First-grade Intelligence Test.* Yonkers-on-Hudson, N. Y., World Book Company, 1928, pp. 5.

The Pribble-McCrory Diagnostic Tests in Practical English Grammar. Chicago: Lyons and Carnahan, 1928.

Allport, Gordon W. and Allport Floyd H.: *A Scale for Measuring Ascendence-submission.* New York: Houghton Mifflin Company, 1928.